DOCUMENT RESUME

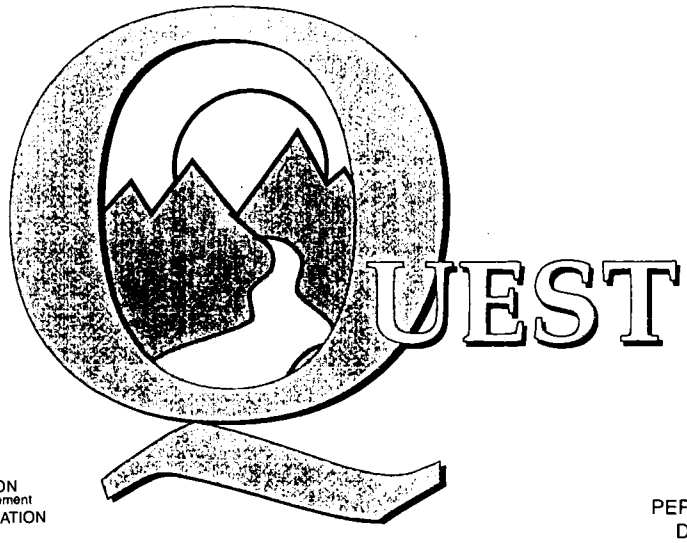
ED 433 358 TM 030 028

AUTHOR Meehan, Merrill L.; Orletsky, Sandra R.; Sattes, Beth
TITLE Field Test of an Instrument Measuring the Concept of Professional Learning Communities in Schools.
INSTITUTION Appalachia Educational Lab., Charleston, WV.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE 1997-07-00
NOTE 74p.
CONTRACT RJ96006001
AVAILABLE FROM Appalachia Educational Laboratory, P.O. Box 1348, Charleston, WV 25325-1348; Tel: 800-624-9210 (Toll Free); e-mail: aelinfo@ael.org; Web site: http://www.ael.org
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Educational Improvement; Elementary Secondary Education; *Field Tests; Learning; *Professional Development; Reliability; *Teacher Attitudes; *Teachers; Test Construction; Validity
IDENTIFIERS Learning Communities

ABSTRACT

A new instrument was designed by S. Hord (1996) to assess globally the maturity of a school's professional staff as a learning community. The Appalachia Educational Laboratory agreed to field test this instrument to determine its reliability and validity and to draw conclusions about its use in educational improvement. The instrument consisted of 17 descriptors of a professional learning community grouped into 5 major dimensions or areas. Faculties of 21 schools (6 elementary, 6 middle and junior high schools, and 9 high schools) participated, for a total of 690 teachers. Some analyses used smaller samples. All five internal consistency reliabilities for the dimension items were in the mid 0.80s, and concurrent validities for the 114 high school teachers scores was 0.75. Factor analysis revealed that a unitary factor of all 17 items accounted for 54 percent of the variance. Results show that the instrument did differentiate the faculties in terms of their maturity as learning communities, and that these differences were evident across elementary, middle/junior high, and high school levels. The instrument appears useful as a screening, filtering, or measuring device to assess the maturity of a school's professional staff. (Contains 16 tables and 22 references.) (SLD)

Field Test of an Instrument Measuring the Concept of Professional Learning Communities in Schools



Merrill L. Meehan
Sandra R. Orletsky
Beth Sattes

Appalachia Educational Laboratory, Inc.
P.O. Box 1348 • Charleston WV 25325 • 800/624-9120

Field Test of an Instrument Measuring the Concept of Professional Learning Communities in Schools

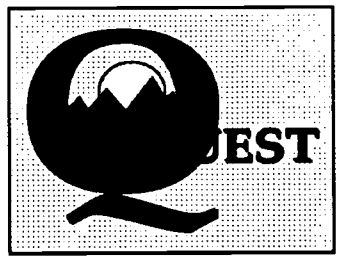
by:

Merrill L. Meehan
Senior Research and Evaluation Specialist

Sandra R. Orletsky
Senior Manager

Beth D. Sattes
Research and Development Specialist

July 1997



QUEST Project
Appalachia Educational Laboratory
Post Office Box 1348
Charleston, West Virginia 25325

AEL's mission is to link the knowledge from research with the wisdom from practice to improve teaching and learning. AEL serves as the Regional Educational Laboratory for Kentucky, Tennessee, Virginia, and West Virginia. For these same four states, it operates both a Regional Technology Consortium and the Eisenhower Regional Consortium for Mathematics and Science Education. In addition, it serves as the Region IV Comprehensive Technical Assistance Center and operates the ERIC Clearinghouse on Rural Education and Small Schools.

Information about AEL projects, programs, and services is available by writing or calling AEL.

Appalachia Educational Laboratory
Post Office Box 1348
Charleston, West Virginia 25325-1348
304/347-0400
800/624-9120 (toll-free)
304/347-0487 (Fax)
aelinfo@ael.org
http://www.ael.org

This publication is based on work sponsored wholly or in part by the Office of Educational Research and Improvement, U. S. Department of Education, under contract number RJ96006001. Its contents do not necessarily reflect the views of OERI, the Department, or any other agency of the U. S. Government.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (continued)

EXECUTIVE SUMMARY

Based on an extensive study of one school that reinvented itself over several years plus a review of educational and organizational literature, Dr. Shirley Hord developed an instrument designed to assess globally the maturity of a school's professional staff as a learning community. As part of its contract to develop a framework for continuous improvement in schools in its four-state region, Appalachia Educational Laboratory (AEL) staff agreed to field test Hord's new instrument. The three objectives of AEL's field test were (1) to assess the instrument's reliability, (2) to assess the instrument's validity, and (3) to draw conclusions about its use in educational improvement efforts. The instrument consisted of 17 "descriptors" of a professional learning community grouped into five major dimensions or areas. The response option for each descriptor was a 5-point scale with different descriptive sentences under the end points and the middle value. Thus, each of the 17 descriptors had three different sentences under its 5-point scale. The reliability was measured by Cronbach's Alpha for internal consistency and by the stability (test-retest) method. Content validity was assessed in its development and reviewing phases. Concurrent validity was assessed through the parallel administration of a school climate instrument. Construct validity was measured two ways: by the "known group" method and by exploratory factor analysis. The sample consisted of the faculties of 21 volunteer schools in the AEL Region—6 elementary, 6 middle/junior high, and 9 high schools. The total number of teachers in the study was 690, although the test-retest reliability and concurrent validity analyses were completed with subsamples of three and four high school faculties in one state, respectively.

In terms of the field-test results, all five internal consistency reliabilities (Alphas) for the dimension items were in the mid .80s and the Alpha for all 17 items was .94. The stability (test-retest) reliability for the 23 high school teachers who could be matched by an identification number was .61. The concurrent validity for 114 high school teachers' scores on the school climate instrument was .75. With respect to the "known group" validity, the higher scores from the teachers in the school known to be a continuous learning community differed significantly (.0001 level) from the teachers in the field test on the five dimensions and the total instrument scale. Factor analysis revealed that a unitary factor of all 17 items accounted for 54% of the variance.

Comparing the Hord instrument results across the 21 field-test schools showed that it did differentiate the faculties in terms of their maturity as learning communities and, further, these differences in maturity were evident across the elementary, middle/junior high, and high school levels. From these results, the AEL researchers concluded that Hord's 17-item instrument is very useful as a screening, filtering, or measuring device to assess the maturity of a school's professional staff as a learning community, especially when the total score is used. Further studies of the stability reliability and the concurrent validity of the instrument in its present form should be completed, especially with more urban schools and with elementary and middle/junior high school faculties. Another recommendation is to include other instruments in further concurrent validity research. If there is interest in measuring the differences of school faculties in the five dimensions of professional learning communities, then it is recommended that the instrument be revised, reformatted and lengthened, and then pilot tested and field tested.

INTRODUCTION

The targets of education reform have shifted dramatically over the past four decades. In the decade of the 1960s, educational reform focused most directly on classroom teachers, both in service and in preparation, through competency/performance-based teacher education efforts and other innovations. In the decade of the 1970s, the targets for reform efforts moved to students through efforts such as minimum competency tests and increased graduation requirements. In the 1980s, the popularity of the effective school research studies and the publication of the provocative book, *A Nation at Risk* (National Commission on Excellence in Education, 1983), moved the reform target to that of the school itself. The effective schools research, especially, featured the importance of the school building principal as the instructional leader of more successful schools—along with several other factors. This shift to the school as the target of reform efforts continues in the 1990s. In fact, the term "restructuring schools" has replaced "reforming schools" and "improving schools" in the literature. Neuman and his associates (1996) write that restructuring schools seems so appealing "...because it suggested that monumental changes were necessary; terms like *improvement*, *innovation*, or *reform* were not robust enough to describe the challenge" (emphasis in original, pg. 5).

Educational reform in the 1990s also experienced a dramatic interest (or renewal of interest, some write) in the concept of community in educational reform efforts. As Plank (1997) noted, there are as many definitions of community as there are authors to write about it. Hord (1996a) notes the term "learning community" is being "...well integrated into the lexicon of American education" (pg. 2). She notes its multiple meanings to various writers. One meaning is that of expanding the traditional classroom into the broader community to use more and different resources and locales to provide improved learning situations and experiences for school children. A second meaning identified by Hord is when members of the larger community come to the school "to enhance the curriculum and learning tasks for students" (pg. 2). A third meaning of community is the full assembly of a school's groups—students, teachers, and administrators—all engaged in the learning process. A fourth meaning of learning community for Hord, and the focus of her current work, is that of a professional community of learners. She defines this meaning of community as "...the teachers in a school along with its administrators continuously seek and share learning, and act on their learning" (pg. 2).

Hord's Conceptualization of Professional Learning Communities

Shirley Hord has been involved in school change and improvement for many years and in a variety of roles. First, she was a faculty member of schools endeavoring to improve continuously. Next, she was heavily involved in the development and dissemination of the Concerns-Based Adoption Model (CBAM). Working with Gene Hall and other researchers at the University of Texas at Austin, numerous instruments, papers, articles, monographs, and books on CBAM were published and utilized in many schools, districts, states, and nations in the world. In her current role, she is an external facilitator of school improvement efforts in five southwestern states and, in that capacity, had the opportunity to work with, study, and document the continuously improving efforts of several schools—one in particular—over a four-year period.

As part of a long-term, qualitative project to help facilitate leadership for change in selected schools in the southwest U.S., Hord worked closely with one elementary school in Louisiana that contributed much to her conceptualization of a professional learning community. This school was close to being shut down due to declining enrollment, but over several years and even with several changes in the principal, the school was able to reinvent itself and flourish. In a series of research reports (Boyd & Hord, 1994a; Boyd & Hord, 1994b; and Hord & Boyd, 1995), the various steps and processes completed by the principals and the professional staff through their reinvention are described and analyzed.

The potential of closure for the school led to the assignment of a new principal who came with the mission for the school focused primarily on the "characteristics and capacities that children brought with them to school" (Hord, 1996a, pg. 17). This focus on the students and adapting the school to fit them was accompanied by the new principal's vision that included staff who became involved in shared decision making and who would be supported by continuous staff development leading to wise decisions. This principal advocated a "person-centered" approach to the school. Through the restructuring of the school's weekly schedule, a special two-hour block of time was reserved for Faculty Study. Faculty Study was a regular discussion and sharing session for the school's professional staff. Led by the principal, staff interacted on various professional development resources obtained, observed, or read about in the past two weeks through conferences attended, other schools visited, or by individual reading. Through this regular sharing and discussion session, a growing vision for the school started to evolve and the faculty grew professionally.

Hord and Boyd report in their research that the second principal continued the successful Faculty Study sessions when he was assigned to the school. But, the researchers observed that this second principal added to the further development of mutual trust and caring relationships among the professional staff. They reported that he helped pull the staff together "in recreational ways" (Boyd & Hord, 1994b, pg. 17) for increased bonding. For example, he suggested staff volleyball games in the gym or group meals at a restaurant to celebrate the accomplishments of the week. Occasionally, potluck suppers with staff's families took place. Thus, the affective dimension of the professional community was enlarged. Not that these social events were the main focus of the second principal, for he did help the staff to identify student learning problems. He then helped the teachers to study and resolve those problems in the regular faculty group sessions.

Then, a third principal came to the school. Hord and her associates reported that her goal was that of enabling students, parents, and staff to work together in more meaningful ways to combine the growing expertise to the greater good of the whole school. To effect her goal, this principal initiated several communications systems within the school and with the parents. For example, she developed one system to announce regular logistical information so it did not occupy valuable time at staff meetings. "She streamlined administrative procedures and organized a management team so that teachers could have an efficient voice in decisions but also spend their time on professional tasks" (Hord, 1996a, pg. 17). She encouraged and supported her teachers in writing proposals to obtain grant money for the school. Also, to help recognize teachers for their unique talents, she encouraged special events.

The fourth principal assigned to the school continued the staff's process of working together, including the weekly Faculty Study session and supporting the culture of the school. To this, he brought a renewed emphasis on curriculum and student learning tasks. Additionally, he brought two other components. The first was computer technology as an instructional tool. The second was the study and subsequent adoption of a curriculum that unified the staff and provided a more consistent curriculum for all students. This latter feature fostered the staff's vision of multiculturalism. This principal continued the staff's reflective dialogue and group dynamics and aided in the staff working together to solve conflicts. To help students resolve their conflicts, a peer mediation program was implemented in the school. Hord and Boyd (1995) describe how the school staff studied, discussed, visited in other schools, and attended a major conference about a new curriculum for the school. The school staff learned about the new curriculum together over the course of a year, then made their decision collectively.

In addition to her extensive experiences in school change research, school improvement processes, and facilitating and studying school leadership, as described in the above school, Hord also draws from contemporary business and corporate literature in the formation of her conceptualization of a professional learning community. She draws on Peter Senge's book *The Fifth Discipline* (1990) for ideas regarding "learning organizations" and how they might apply in educational organizations such as schools. She quotes Senge's notion of a learning organization "where new and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning how to learn together" (Senge, 1990, pg. 3). Senge wrote that a learning organization focused on five main disciplines: shared vision, mental models, team learning, personal mastery, and systems thinking. Hord identifies Senge's concept of organization design as a radical departure from traditional thinking and as strongly influencing educational theorists and writers. Hord (1996a) states that when writers applied Senge's concept to schools, the term "learning communities" was used (page 6).

Hord (1996a) also draws on other writers in the business and corporate literature base for inputs to her formation of a professional learning community. She cites Deal and Kennedy (1982) who reported that business and industry leaders used cultural factors to cause changes in their staff. To Senge's seminal book, Hord adds Block (1993), Galagan (1994), and Whyte (1994) as contributing to a "stream of writing [that] emphasizes the importance of nurturing and celebrating the work of the individual staff person, and of supporting the collective engagement of staff in such activities as shared vision, development, problem identification, learning, and problem resolution" (Hord, 1996a, pg. 3).

Hord (1996a) continues her review of the educational and business literature, organized by key attributes that she feels comprise the major components of professional learning communities. Hord's five major attributes of a professional learning community are supportive and shared leadership, collective creativity, shared values and vision, supportive conditions, and shared personal practice.

Development of the Hord Instrument

Hord's current project for the Southwest Educational Development Laboratory (SEDL) continues her work as a facilitator of change and improvement at the school level in the five-state SEDL region. She identified a "context conducive to change" as an essential ingredient for successful school improvement. Further, her research showed professional learning community to be a key factor in the context of improving schools. Hord's present work will explore the concept of professional learning community by studying several schools whose staff function as learning communities, identifying components of how they operate as learning organizations, and then developing strategies and materials to help other schools become professional learning communities.

After her review of the educational and business literature, Hord's next step was to identify several schools to study how they operate as learning communities. To complete this step, she needed a small set of schools. Nominations of such schools would be sought from contact persons in SEDL's states, but Hord had to identify the criteria for nominating schools. To do this, she developed an instrument to serve as a screening or filtering tool. When the contact persons nominated a school for Hord to study as a professional learning community, they would complete her new instrument and send it back to her. In this way, she would then have a set of completed instruments to compare the schools on the criteria she developed.

Hord developed the materials for contact persons to nominate schools as learning communities in the early summer of 1996. These materials consisted of two coordinated items: a cover letter explaining her request and the instrument containing her criteria for professional learning communities, which included a form soliciting descriptive information about the nominated school and the contact person's relationship to it. This study is about the latter item.

The instrument developed by Hord (1996a) was titled "Descriptors of Professional Learning Communities." It consisted of 17 descriptors grouped by five major attributes or areas identified from prior research and a review of the literature. These five attributes or areas are:

- the collegial and facilitative participation of the principal who shares leadership and, thus, power and authority through inviting shared decision making from the staff;
- a shared vision that is developed from the staff's unswerving commitment to students' learning and that is consistently articulated and referenced for the staff's work;
- learning that is done collectively and work that applies the learning to create solutions that address students' needs;
- the visitation and review of each teacher's classroom by peers as a feedback and assistance activity to support individual and community improvement; and
- physical conditions and human capacities that support such an operation (p. 10).

The format of the Hord professional learning community instrument was unique. As stated above, the 17 descriptors were grouped into five major attributes or areas. These major areas were designated by Roman numerals. The descriptors were under the short explanations of the major attribute, generally, as stated in the previous paragraph. None of these descriptors were numbered. Each descriptor consisted of a 5-point scale, from "5" (high) to "1" (low), inside a narrow rectangular box spreading across most of the 8½" width of the page. Immediately beneath each 5-point scale box were three indicators or sentences to differentiate the high, middle, and low points on the scale. That is, the first indicator fit under the "5" option on the scale, the second indicator fit under the "3" option, and the third indicator fit under the "1" option. Each indicator or sentence spread out a little beyond the width of the single number, but it was clear that these indicators represented the high, middle, and low response options. For example, the three indicators under the 5-point response scale for the second descriptor, which was the last one under the first major attribute or area, from left to right were:

- Administrator(s) involves the entire staff.
- Administrator(s) involves a small committee, council, or team of staff.
- Administrator(s) does not involve any staff. (Hord, 1996b)

The "Descriptors of Professional Learning Communities" were printed on four 8½"x11" pages in the vertical format. There were no provisions for the school's identifying information, as this was elicited on the cover page form described above. The directions for the instrument asked the contact person to "...consider where you believe the nominated school is in its development of each of the five Roman-numbered descriptors (below)" (Hord, 1996b, pg 1). The directions then asked the respondent to mark his/her assessment on the 5-point scale above the three indicator statements. The respondents were asked to mark any one of the five numbers in the box of each response scale, and not in between the numbers by inference. Thus, the unique format and layout of Hord's new instrument on professional learning community required the respondent to read all three indicators for each of the 17 descriptors and then mark the response scale. This format and layout required more mental processing than usual for an instrument, but this was by design because of its intended use as a screening or filtering device. A copy of Hord's new instrument, as developed in the summer of 1996, appears as Appendix A in this report.

Background and Objectives of This Study

The Appalachia Educational Laboratory (AEL) is a regional research and development organization (like SEDL) serving the states of Kentucky, Tennessee, Virginia, and West Virginia. In its latest five-year contract, AEL designed one regional project, called QUEST, to assist schools in their educational reform efforts. QUEST is an inquiry-based journey of systemic transformation designed to challenge the norms embedded in the culture of traditional schools (Appalachia Educational Laboratory, 1995). QUEST proposes to (a) establish a network and (b) develop a process, both of which will harness the power of collective thinking and collegial learning for continuous improvement in schools. The two goals of QUEST are (1) to produce a framework and a process that will enable members of school communities to embark on a journey for continuous

improvement and (2) create a network of individuals and agencies to sustain and support those school communities in their journey. AEL will develop or adapt several tools, techniques, and processes to assist groups in action research, reflection, and improved understanding of complex improvement issues through collaborative study and disciplined discussion. Examples of such techniques to be employed by AEL staff include collegial investigations, interview design process, and story investigation (Appalachia Educational Laboratory, 1995, Section II).

As a research and development organization, AEL is committed to evaluating its efforts, including the QUEST regional project. The evaluation of the QUEST project led to a search for evaluation instruments with potential use for judging its implementation and outcomes. The unique nature of the evolving processes of QUEST's continuous journey for school improvement meant that there may not be any evaluation instrument exactly matching the QUEST effort and that, ultimately, a new instrument may have to be developed by AEL staff. Nevertheless, AEL staff located several current instruments, including Hord's new one, that measured several of the components in the QUEST project. Most of the instruments had empirical data regarding their psychometric properties. However, being brand new, the Hord instrument had not been tested in the field with a sample of schools, nor were there plans to do so, given the limited resources and staff that SEDL allocated to their project. AEL staff offered to pilot test and field test Hord's new instrument on professional learning community to obtain and report psychometric data for AEL's possible use and for SEDL's use, too.

AEL staff pilot tested Hord's professional learning community instrument in the first QUEST conference/retreat for school teams in the summer of 1996. The instrument was reformatted and retitled for the pilot test. (These changes by AEL staff are explained in the next major section of the report.) Two other potential evaluation instruments were pilot tested at the same time. A total of 28 QUEST conference/retreat participants, representing 11 West Virginia schools, completed the set of instruments before the first official QUEST retreat activity. Later, in September, each participant was asked to complete the Hord instrument only and mail it back to AEL for computing the test-retest reliability figure. The internal consistency reliability was assessed by Cronbach's Alpha and concurrent validity was measured by computing the Pearson product-moment correlation between the total scores on the Hord instrument and a school climate instrument (Manning, Curtis, and McMillen, 1996).

The pilot test of the new instrument, renamed "School as Learning Organization" by AEL staff, with the small group of West Virginia students, parents, and educators participating in the summer 1996 QUEST project conference/retreat was very positive. The Alpha reliabilities for the items in the five major areas were +.84, +.68, +.82, +.78, and +.83 in order, while the Alpha reliability for the total of 17 items was +.92 in the pilot test. The test-retest reliabilities for the 15 participants who could be matched with individual ID numbers were +.94, +.86, +.73, +.86, and +.78 for the items in the five major areas and +.94 for all 17 items together. The correlation of the "School as Learning Organization" instrument total score to the total score of the school climate instrument was +.82 (Appalachia Educational Laboratory, 1997). The pilot test of the Hord instrument in the AEL Region with a small, heterogenous group proved that it showed the promise of possessing the

psychometric properties sufficient to continue its use, but a field test with a larger sample of schools was required.

The major purpose of this study was to field test Hord's professional learning community instrument in schools in AEL's four states to discover its psychometric properties and assess its potential for use in educational improvement projects at the school building level. A related purpose was to document and report on the instrument's psychometric properties for other educational researchers and evaluators working in this field. More specifically, the three objectives of this study were:

- to assess the reliability of Hord's instrument on professional learning communities,
- to assess the validity of Hord's instrument on professional learning communities, and
- to draw conclusions about its use in educational improvement efforts at the school level.

Audience for This Report

The primary audience for this report of the field test of Hord's instrument on professional learning communities is the staff of AEL's QUEST project. This report is seen as providing important new information about an instrument that has possibilities for use in the evaluation of the QUEST project. This field-test report presents usability, reliability, and validity information for QUEST staff's decision-making regarding the Hord instrument. Secondary audiences include Shirley Hord at SEDL and other researchers and evaluators interested in measuring the degree to which a school functions as a professional learning community (as defined by Hord) and the processes, outcomes, and impacts of school building-level improvement projects, especially with the professional staff.

METHODOLOGY

This section presents descriptions of the sample and subsamples in the field test, a description of how AEL staff modified the instrument used to collect the data, and the methods for collecting the field-test data and the analyses employed.

Sample and Subsample Descriptions

The sample for this study included all the teachers in 21 schools in AEL's four-state Region who completed and returned the "School as Learning Organization" instrument. The 21 schools in this field test volunteered to participate in the study with no external rewards or motivation offers. These schools were "nominated" to participate in this effort, usually through the building principal or by another contact person familiar with the school and its staff. A total of 690 teachers in the AEL Region completed and returned the instrument on professional learning communities. Since the school building is the target for this instrument, no demographic data were sought from the teachers. This cut down on administration time as well as assisting in gaining agreement to cooperate in the study.

Some selected information about the 21 schools in the study was obtained from data in state school guides (Quality Education Data, 1996a, 1996b, 1996c, & 1996d). The field test schools were in the following states: four in Kentucky, six in Tennessee, four in Virginia, and seven in West Virginia. With respect to school levels, six were at the elementary level, six were at the middle/junior high level, and nine were at the high school level. The enrollment of students in the field-test schools ranged from a low of 205 to a high of 1,200 with the median at 568 students enrolled. The mean enrollment figure was 612 with a large standard deviation of 294 students. Information on the number of students on free and reduced lunches was available at the ***district*** level for the 21 schools. The percent of students on free and reduced lunches in the 21 districts ranged from a low of 12% to a high of 39%, with a mean of 22.48% and a standard deviation of 9.45%.

A subsample of teachers in four large high schools in Tennessee volunteered to participate in the concurrent validity and stability (test-retest) reliability analyses by (1) completing a school climate instrument at the same time and (2) by including an individual identification number on their instruments. Per past successful experiences with identification numbers, the teachers in these four Tennessee high schools were asked to write the last four digits of their Social Security number on the top right corner of the professional learning community instrument (at both administrations) and on the school climate instrument.

The subsample high schools (grades 9-12) were large with enrollments of 1,200; 868; 1,015; and 1,099 students. The number of teachers at each high school completing the first administration of the instrument was 53, 57, 61, and 60 (in the order of the enrollments in the prior sentence). These schools are the only high schools in the county school system. The district also has 19 elementary schools and 8 middle schools. Located in eastern Tennessee, this district's student population is 99%

Caucasian, with 13% on free or reduced lunches. It is reported that 64% of these high school students are college bound, based on the percentage of last year's graduating class that enrolled in two- or four-year colleges (Quality Education Data, 1996b, pg. 38).

Hord's Professional Learning Community Instrument—Reformatted

AEL QUEST project staff revised Hord's original professional learning community instrument (see Appendix A) slightly for this field test. AEL's revision was not very extensive, being mainly cosmetic in nature. Further, Hord was apprised of AEL's minor revisions and approved of them for the field test. These minor revisions are described in the following paragraphs.

The first step AEL staff completed with Hord's 17-item instrument was to review the descriptor and indicator statements for consistency. AEL staff did not attempt to change the intent or language of any of the descriptors or indicators; rather, they checked for consistency of language across these two parts. Where necessary, slight adjustments or word changes were made in an attempt to increase consistency, not alter Hord's original meaning and intent.

The second step completed by AEL staff was to reformat and typeset the instrument. In its original form, the instrument consisted of 17 items (called descriptors) grouped under five Roman-numeral areas. These elements were typed on the pages in the vertical format; requiring four pages for the 17 items. Demographic data were elicited by a separate form that was on top of the four pages of items (descriptors). Respondents were asked to read the trio of indicators under the rectangular response box with the numbers of "5" through "1" inside, then circle or otherwise mark one of those five numbers. The three indicator statements helped the respondent choose which number to mark for each item (descriptor) because they reflected the high ("5"), middle ("3"), and low ("1") positions on the response continuum. Again, the reader is referred to Appendix A for the original version of the Hord instrument.

In reformatting the original instrument, AEL staff felt it could reduce the number of pages for the 17 items by switching to a horizontal format. Too, AEL staff felt it would improve the usability of the instrument if the five major area statements were set off to the left of the 17 individual items. Also, the numbering system was changed to be single digit Arabic numbers for the five major area statements and those numbers plus lower case letters for each of the items associated with a major area (e.g., 1a, 1b, 2a, 2b, 2c, etc.). The 5-point response scale for each item was retained, along with those five numbers, but a single horizontal line with short, vertical point markers replaced the original rectangular box. Two additional graphic additions included (1) a bracket at the left side of the response line and indicator statements and (2) a thin, horizontal broken line between each of the five major areas. These two additions were made to help respondents see which area they were responding to. These reformatting changes can be seen on AEL's version of the instrument appearing as Appendix B to this report.

Four other reformatting changes were made by AEL staff—all to the top quarter of the first page. First, to reflect how the instrument might be used in the QUEST project, the title was changed

to "School as Learning Organization." Second and third, lines for the respondent to write in the date and the school name were provided. Since an identification number for each participant in the pilot test was required and, presumably, also would be required in any projects **other** than the field test, the version in Appendix B solicits the last four digits of the respondent's Social Security number. Fourth, the directions were rewritten and reformatted for AEL's use, since it was not designed to be administered with a formal cover letter like the original version (see Appendix B).

In sum, AEL staff worked very hard to reformat Hord's original instrument to be shorter and more user-friendly, yet not change any of its original intents or meanings. AEL's redesigned version of Hord's instrument (Appendix B) used three pages in the horizontal format to present the 17 items in a coordinated, visually appealing manner.

Study Procedures

The administration of the "School as Learning Organization" instrument was conducted between November 1996 and May 1997 for this field test. These seven months included the single administration of the instrument to most schools, the administration of the "School as Learning Organization" instrument and the school climate instrument to a subsample of high schools, and the readministration of the target instrument to three of the four high schools in the subsample. All 21 schools volunteered (or were volunteered by someone) to complete the instrument during school time.

Schools participating in this field test were suggested by contact persons in AEL's four states. The AEL QUEST staff solicited schools to participate in this study from educators and member AEL Board of Directors. Several of the schools worked with AEL in past projects, but just as many were unknown to AEL. There were no rewards or incentives for schools to participate except that a summary report of the school's analyzed data would be sent to both the contact person and the school building principal (if they were different individuals).

Once staff from a contacted school agreed to participate in the field test, they were asked to call AEL directly to provide their name, mailing address, and how many copies they needed. AEL support staff then copied, boxed, and shipped the blank instruments to the building-level person and included a pre-addressed shipping label for returning the completed instruments. The field test version of the reformatted instrument did not include room for the respondent's last four Social Security number digits, as matching was not a crucial step, except for the concurrent validity and stability (test-retest) reliability analyses.

Completed sets of the "School as Learning Organization" instruments were returned to AEL starting in December 1996 and continued up until May 1997. As the packages of instruments were received at AEL, they were checked to make sure their identification information was complete. This information was entered into a "running" file of schools. As more schools' packages were received, they were added to the database of cooperating schools.

Next, the sets of completed instruments were sent to a different work unit within AEL where the instruments were entered into an SPSS database. Since each school was promised an individual summary report, the school identification codes were used in analyzing the data after it was entered in the main database. The data for each school was analyzed soon after it was entered. Then, an individual summary report was made for each school from the SPSS printouts. These summary reports included the numbers, means, and standard deviations for all 17 items. Also, an "X" representing the mean was hand drawn on each item's response line to illustrate graphically where the school faculty's average response rating was located on the 5-point scale. Also, the five major dimension numbers, aggregated item means, and standard deviations were provided in the open space under each dimension's descriptive sentence on the left side of the page. Finally, the overall Alpha reliability coefficient for the instrument was typed on the upper right corner of the first page.

The procedures for the subsample differed from the single-administration schools described above. The subsample consisted of four large high schools in a county school district in Tennessee. The contact person for these four schools was an AEL Board member who volunteered to participate more fully than others when approached by the AEL staff. It was decided that this contact person could help with the concurrent validity and stability (test-retest) analyses of the field test because she offered to ask the school administrators if they would present it to the faculty to vote on. All four high schools voted to participate in the additional instrument administrations, provided they received summary reports of the concurrent validity school climate instrument. AEL agreed to provide these additional summary reports along with the summary reports for the "School as Learning Organization" instrument.

Finally, in addition to the 21 schools in the AEL Region in the field test, the instrument was administered to a school known from prior research to be in a continuously-improving mode (S. M. Hord, personal communications, May 1997). This elementary school was one of the schools studied by Hord and her associates for several years. In fact, this special school in SEDL's region was the inspiration for some of the five major dimensions in the instrument and many of the items written by Hord. This school, a "known group" for the construct validity analysis, is an urban school of about 400 students in the New Orleans school district. Hord administered the instrument for AEL as part of this field test (S. M. Hord, personal communications, May 1997). Nineteen instruments were sent to AEL for the "known group" analysis but, as usual, not every teacher completed every item.

Data Analyses

The analyses of the "School as Learning Organization" instrument began with the main file of 690 teachers; however, several analyses required the construction of separate files of data from the four high school faculties in the subsample. Also, the file of teachers from the "known group" school was added to the main file for one other analysis. The analyses of these files are presented below in paragraphs describing the descriptive statistics, the reliability analyses, and the validity analyses. A much smaller file of information about the 21 schools in the field test was used to describe the sample and subsamples above and is not discussed in this section. All of the analyses were completed using the SPSS statistical analysis software package at AEL in Charleston, West Virginia.

Descriptive analysis of the 690-case file was the first step completed. All of the descriptive statistics for the 17 individual instrument items were computed, followed by the descriptive statistics for those grouped by the five major dimensions of the instrument. Next, those same descriptive statistics were computed by the school level—elementary, middle/junior high, and high school. Then, as one measure of the usability of the instrument, these same descriptive statistics (individual items and dimensions) were computed for the 21 different schools in the field test.

Reliability analyses consisted of two types—internal consistency and stability (or test-retest). First, the internal consistency reliability coefficient, using Cronbach's Alpha formula, was computed for each of the five major dimensions and the total instrument. These Alpha reliability coefficients were computed on the main file of 690 cases although, as usual, not all teachers completed all the items, resulting in varying numbers of cases for the six Alphas. Next, the instrument's Alpha reliabilities for the five dimensions and the total instrument score were computed for the 21 individual schools in the field test. These analyses were conducted to assess the reliabilities at the level of intended use—the individual school. Second, the stability (test-retest) reliability coefficient was computed with a subsample of three high school faculties in Tennessee. Recall that none of the instruments used in the field test requested the respondent's last four Social Security number digits. Rather, the teachers in the subsample were asked by the instrument's administrator to insert these digits on both instruments at the initial administration and at the follow-up administration. Not all teachers volunteered to provide this unique identification number either initially or later; or they forgot or switched numbers. This caused the number of cases that could be matched in a new combined file to drop dramatically to just 30, and only 23 of those cases had a viable total instrument score. Although this was less than an ideal situation for assessing the stability reliability, there was no viable alternative in this study.

Validity analyses consisted of three types—content, the "known group" method, and factor analysis. First, content validity was assessed in the reviewing and reformatting stages. Second, for the "known group" construct validity analysis, the scores of the teachers in the school known from prior research to be functioning as a professional learning community were compared to the 690 teachers from the 21 schools in the field-test database. The 21 AEL schools were volunteer schools and it was not assumed that they were or were not schools of continuous learning and improvement, and there were no data available to either support or refute that. The purpose of this second validity check was to assess the difference of the scores from the "known group" teachers with the scores from all the other teachers in the main database with the t-test. The employment of the t-test in this analysis violates several of the assumptions for its use and, although this test is very robust to such violations, the SPSS software analyzes and adjusts for such violations, as appropriate.

Specifically, the SPSS software automatically computes the Levene test that the two samples came from populations with the same variances. As expected, all six Levene tests were significant, showing that the two variances were not equal. In these cases, the SPSS t-test printout for the "unequal" row is read where degrees of freedom for the t-statistic are calculated based on both the sample sizes and the standard deviations of each of the two groups. Thus, the computation of the t-statistic is adjusted for the unequal variances in the two samples and this adjusted t-value, degrees of freedom, and other values are included in the SPSS printout. Third, factor analyses included

unconstrained principal components analysis followed by both varimax and oblique rotations of the data. The final factor analysis solution was an interactive process of studying the before-rotation data with the after-rotation data, then going back to the descriptive statistics on the scores, including their distributions.

FINDINGS

This section presents the findings from the field test of Hord's new instrument for measuring professional learning communities in schools. These findings are presented in tables and narrative copy organized by the major topics of descriptive statistics, reliability results, and validity results, with subsections as appropriate.

Descriptive Statistics and Usability

Full Group Statistics

Table 1 displays the descriptive statistics for the 17 individual items in the Hord instrument by the full group of teachers in the field test. Not all teachers responded to every item in the instrument, which is normal. The second column shows that the number responding to the 17 items ranged from the low of 649 (on item 4b) to the high of 688 (on items 2b and 4a). The data in the minimum and maximum score columns show that the teachers used the full range of possible scores (1.00 to 5.00) on all 17 items. The fifth column reveals the mode responses as 3.00 or 4.00 for 15 items and 2.00 for items 4a and 4b. All of the median responses were also 3.00 or 4.00, except for item 4a, which was 2.00. The means for 15 of the 17 items ranged from 3.29 to 3.85 on the 5-point scale, except for items 4a and 4b, which were 2.32 and 2.75, respectively. The standard deviations for all items were in the range of 0.91 to 1.20. The coefficient of variation values, which are computed by dividing the standard deviation by the mean score, ranged from .239 up to .433, with the values for items 4a and 4b, being almost a tenth larger than the largest of the other values. For the standard error values, 12 of the 17 were in the .30s and 5 were in the .40s. The last column displays the skewness value for each item and the deviation—positive or negative—from a perfectly skewed distribution. Data in this last column show just 3 items were positively skewed and the remaining 14 items were negatively skewed. Items 1a, 1b, 4b, and 5b were the closest to a normally-skewed distribution, while items 3e, 5c, 5a, and 2b were the farthest from a normal distribution.

Table 2 displays the descriptive statistics for the five dimensions and the total instrument score by the full group of teachers in the field test. For the reader's convenience, shortened titles, drawn from the five major dimensions on the instrument, appear in the first column. These dimension names will be used consistently throughout the remainder of this report. Column two shows the number of respondents with dimension scores varied with 649 to 679 and that 595 respondents completed all 17 items to yield a total instrument score. The number of items per dimension varied with 2 (dimensions #1 and #4), 3 (dimension #2), and 5 (dimensions #3 and #5). Columns three and four reveal that the teachers used the full range of available responses for each dimension and the total score. Column five shows the mode responses were above the midpoint on all dimensions except #4 and the total score, while column six shows the median scores were slightly above or below the mode response but always at or above the midpoint value for the five dimensions and the total score. And, the mean scores were all above the mid-points, with dimension #4 being the closest to the mid-point value. The standard deviations varied, thus the coefficient of variation column is a better indicator

Table 1

Descriptive Statistics for the 17 Individual Items
by the Full Group of Teachers

Item Number	Number of Teachers	Minimum Score Obtained	Maximum Score Obtained	Mode	Median	Mean	Standard Deviation	Coeff. of Variation	Standard Error	Skew-ness
1a	685	1.00	5.00	3.00	3.00	3.32	1.05	.316	.040	-.030
1b	680	1.00	5.00	3.00	3.00	3.41	0.93	.273	.036	.043
2a	686	1.00	5.00	3.00	3.00	3.49	0.99	.284	.038	-.325
2b	688	1.00	5.00	4.00	4.00	3.85	0.98	.255	.037	-.495
2c	683	1.00	5.00	4.00	4.00	3.62	0.98	.271	.038	-.344
3a	685	1.00	5.00	3.00	3.00	3.35	1.06	.316	.041	-.369
3b	683	1.00	5.00	3.00	3.00	3.37	1.03	.306	.039	-.164
3c	684	1.00	5.00	4.00	4.00	3.57	0.95	.266	.036	-.485
3d	684	1.00	5.00	4.00	4.00	3.80	0.91	.239	.035	-.436
3e	686	1.00	5.00	4.00	4.00	3.65	0.94	.258	.036	-.565
4a	688	1.00	5.00	2.00	2.00	2.32	1.00	.431	.038	.419
4b	649	1.00	5.00	2.00	3.00	2.75	1.19	.433	.047	.067
5a	662	1.00	5.00	4.00	4.00	3.54	1.20	.339	.046	-.539
5b	666	1.00	5.00	3.00	3.00	3.29	1.08	.328	.042	-.076
5c	667	1.00	5.00	4.00	4.00	3.49	1.01	.289	.039	-.560
5d	669	1.00	5.00	3.00	3.00	3.39	0.97	.286	.038	-.168
5e	667	1.00	5.00	4.00	4.00	3.55	1.01	.285	.039	-.397

Table 2

Descriptive Statistics for the Five Dimensions and Total Score by the Full Group of Teachers

Dimension Number and Name	Number of Teachers	Minimum Score Obtained	Maximum Score Obtained	Mode	Median	Mean	Std. Dev.	Coeff. of Variation	Std. Error	Skew-ness
1. Principal's Facilitative Leadership (2 items)	677	2.00	10.00	6.00	6.00	6.72	1.87	.278	.072	.017
2. Shared Visions for Improvement (3 items)	679	3.00	15.00	9.00	11.00	10.97	2.55	.232	.098	-.286
3. Collective Creativity and Learning (5 items)	673	5.00	25.00	20.00	18.00	17.77	3.90	.219	.150	-.397
4. Classroom Observations and Feedback (2 items)	649	2.00	10.00	4.00	5.00	5.12	2.04	.398	.080	.105
5. School Conditions and Capacities (5 items)	655	5.00	25.00	18.00	18.00	17.25	4.18	.242	.163	-.272
Total Instrument Score (17 items)	595	17.00	85.00	61.00	58.00	57.97	12.33	.213	.505	-.255

of how much the dimensions varied relative to their mean score. Here, we see that five of the six values were in the .200s, with the exception being dimension #4 at .398. Three of the standard errors were under .10, two (dimensions #3 and #5) were .150 and .163 respectively, while the total score standard error was .505. In terms of skewness, four of the six were negatively skewed, while dimensions #1 and #4 were positively skewed. Dimension #1, at 0.17, was very close to being a normally distributed set of scores.

Schools by Levels

Tables 3 and 4 display the descriptive statistics for the five dimensions and the total score for the six elementary schools in the field test—the first four dimensions in Table 3 and the fifth dimension and the total score in Table 4. The number of teachers responding in the six elementary schools ranged from the low of 8 to the high of 40; three schools had 17 or fewer responding teachers, while the other three had 27 or more teachers. Looking at dimension #1, Principal's Facilitative Leadership, we see that school #3 had the highest possible score for the two items—every one of the school's 17 teachers rated both items at 5.00 on the 5-point maximum scale. Thus, there was no deviation, coefficient of variation, standard error, or skewness for school #3 on the first dimension. In contrast, school #6 had the lowest maximum score, the lowest mean score (6.50), the third lowest standard deviation, the third lowest standard error, and the third largest skewness value on the first dimension. For dimension #2, Shared Visions for Improvement, the six schools had very similar minimum and maximum scores, but varied more on the mode and median scores where school #2 had the lowest values (11.00). School #2 also had the lowest mean of 11.00, while the other five ranged from 12.08 to 13.88. School #6 had the largest standard deviation, coefficient value, and standard error value. School #2 had the smallest standard error (.299) and also the smallest skewness value (.264), while school #3 had the largest skewness value (-2.049).

Regarding dimension #3, Collective Creativity and Learning, data in Table 3 show minimum scores varied more (9.00 to 18.00) than the maximum scores (23.00 to 25.00) on the 25-point maximum score. However, both the mode and median scores varied less so (18.00 to 25.00). The mean scores ranged from the low of 17.70 for school #2 up to 23.29 for school #3. The standard deviations ranged from 2.17 for school #3 to 3.46 for school #2. The coefficient of variation values all were less than .196 and school #3 had the smallest at .093. Four of the six standard errors were in the .500s, one was in the .600s, and the largest was .866 for school #5. Five of the six skewness values were negative, and two of those were more than 1.000. The smallest skewness value was -.512 for school #2. On the 10-point maximum score for dimension #4, the minimum scores were either 2.00 or 4.00 for all six schools, while the maximum scores ranged from 8.00 (schools #1 and #5) to 10.00 (schools #2, #3, and #6). Three schools had a mode of 4.00 and the other three were 5.00, 6.00, and 7.00. Their medians ranged from 4.00 to 7.00, similar to their mean scores. The lowest mean was 4.79 for school #4 and the highest was school #3 at 7.06. The standard deviations for five schools ranged from 1.41 to 1.95; school #6 was 2.20. The coefficient of variation values for four schools were in the .300s; for schools #1 and #3 were .238 and .276, respectively. The standard errors ranged from the low of .265 for school #2 up to .627 for school #5. In contrast to

Table 3

Descriptive Statistics for the First Four Dimension Scores
by the Six Elementary Schools

School Number	Number of Teachers	Minimum Score Obtained	Maximum Score Obtained	Mode	Median	Mean	Standard Deviation	Coeff. of Variation	Standard Error	Skew-ness
1. Principal's Facilitative Leadership										
1	29	4.00	10.00	7.00	7.00	7.00	1.51	.216	.281	-.200
2	39	4.00	10.00	6.00	6.00	6.64	1.18	.178	.189	.756
3	17	10.00	10.00	10.00	10.00	10.00	0.00	.000	.000	.000
4	31	4.00	10.00	7.00	7.00	6.71	1.83	.273	.329	.147
5	9	7.00	10.00	10.00	10.00	9.22	1.09	.118	.364	-1.289
6	14	4.00	9.00	6.00	6.00	6.50	1.35	.208	.359	.222
2. Shared Visions for Improvement										
1	29	8.00	15.00	13.00	13.00	12.17	1.85	.152	.344	-.705
2	39	8.00	15.00	11.00	11.00	11.00	1.87	.170	.299	.264
3	17	8.00	15.00	15.00	15.00	13.88	1.93	.130	.469	-2.049
4	30	8.00	15.00	15.00	13.00	12.97	1.87	.144	.341	-.699
5	9	9.00	15.00	14.00	14.00	13.22	1.86	.141	.619	-1.621
6	13	8.00	15.00	13.00	13.00	12.08	2.36	.195	.655	-.871
3. Collective Creativity and Learning										
1	28	11.00	25.00	20.00	20.00	20.00	3.21	.161	.606	-.898
2	40	9.00	24.00	18.00	18.00	17.70	3.46	.195	.547	-.512
3	17	18.00	25.00	25.00	24.00	23.29	2.17	.093	.527	-1.175
4	31	11.00	25.00	20.00	20.00	19.39	2.93	.151	.526	-1.039
5	9	18.00	25.00	19.00	19.00	20.67	2.60	.126	.866	.855
6	14	15.00	23.00	20.00	20.00	20.14	2.21	.110	.592	-.706
4. Classroom Observations and Feedback										
1	29	2.00	8.00	7.00	6.00	5.93	1.41	.238	.267	-1.054
2	39	2.00	10.00	6.00	5.00	5.21	1.66	.319	.265	.058
3	16	4.00	10.00	5.00	7.00	7.06	1.95	.276	.487	.271
4	29	2.00	9.00	4.00	4.00	4.79	1.88	.392	.349	.252
5	8	4.00	8.00	4.00	5.00	5.50	1.77	.322	.627	.615
6	14	2.00	10.00	4.00	6.50	6.07	2.20	.362	.588	-.156

Table 4

Descriptive Statistics for the Fifth Dimension Score and the Total Instrument Score by the Six Elementary Schools

School Number	Number of Teachers	Minimum Score Obtained	Maximum Score Obtained	Mode	Median	Mean	Standard Deviation	Coeff. of Variation	Standard Error	Skew-ness
5. School Conditions and Capacities										
1	29	12.00	25.00	20.00	21.00	20.69	2.63	.127	.489	-1.036
2	39	8.00	21.00	18.00	17.00	16.05	3.11	.194	.498	-.724
3	16	17.00	25.00	25.00	23.00	22.19	2.66	.121	.666	-.472
4	30	11.00	24.00	18.00	18.00	18.13	3.19	.176	.583	-.068
5	9	13.00	25.00	21.00	21.00	20.56	3.43	.167	1.144	-1.350
6	14	16.00	23.00	20.00	20.50	20.50	2.18	.106	.581	-.654
Total Instrument Score										
1	27	52.00	79.00	65.00	66.00	66.37	7.12	.107	1.369	-.269
2	37	37.00	73.00	61.00	58.00	56.81	8.79	.155	1.446	-.531
3	15	68.00	85.00	85.00	79.00	77.87	5.66	.073	1.460	-.313
4	27	45.00	81.00	56.00	61.00	61.41	7.89	.128	1.581	.260
5	8	56.00	78.00	78.00	69.50	69.50	8.02	.115	2.835	-.435
6	13	50.00	77.00	65.00	65.00	65.15	7.88	.121	2.186	-.343

the prior dimension, four of the skewness values were positive and low values, and two were negative values. The skewness value for school #2 was very close to normal at .058.

For dimension #5 in Table 4, the six minimum scores ranged from 8.00 to 17.00 on the 25-point maximum score. The maximum scores ranged from 21.00 to 25.00. The mode scores ranged from 18.00 to 25 for school #3. Their median scores were from the low of 17.00 for school #2 to 23.00 for school #3. School #3 also had the highest mean score at 22.19 and the third lowest standard deviation at .266. By contrast, school #2 had the lowest mean at 16.05 and the fourth largest standard deviation at 3.11. The coefficient of variation values ranged from .106 to .194, while the standard errors varied much more with a low of .498 and a high of 1.144 for school #5 (the school with the largest standard deviation and the smallest faculty). All six skewness values were negative, ranging from -.068 for school #4 to -1.350 for school #5.

Table 4 also displays the total instrument statistics for the six elementary schools. Possible scores on the total instrument ranged from 17 to 85 points. The minimum scores for the six schools ranged from 37.00 to 68.00 (school #3), while maximum scores ranged from 73.00 to 85.00 (school #3). The modes ranged from 56.00 for school #4 up to 85.00 (the maximum attainable) for school #3. The median scores ranged from a low of 61.00 for school #4 up to 79.00 for school #3. The mean scores and the standard deviations showed rather large differences across the six elementary schools. For example, with a mean of 77.87 and a standard deviation of 5.66, school #3 clearly scored much higher on the total instrument than the other five schools. These values resulted in the smallest coefficient of variation value at .073. Four schools had means in the 60s and standard deviations in the 7.00s except one being 8.02. School #2 had the lowest mean score at 56.81 and the largest standard deviation at 8.79, resulting in the largest coefficient of variation value at .155. Thus, the range value for the mean scores of schools #3 and #2 was 21.06 on the 85-point maximum score. The standard error values ranged from the low of 1.369 for school #1 to 2.835 for school #5, with school #6 at 2.186 having the only other value above 2.000 (most likely due to their small faculty sizes). Five of the six skewness values were negative and under .532, with only school #4 positive at .260. None were near normalcy in terms of skewness of distribution of scores.

Tables 5 and 6 display the descriptive statistics for the five dimensions and the total score for the six middle/junior high schools in the field test. Similar to the arrangement above, the first four dimensions are in Table 5 and the fifth dimension and the total score are in Table 6. The number of responding teachers in the six middle/junior high schools ranged from the low of 14 to the high of 52, with three schools having 21 or fewer teachers. Dimension #1 consisted of two items with a maximum score of 10 points. The minimum scores ranged from 3.00 to 5.00, while the maximum scores were all the same at 10.00. Three schools had modes of 6.00, one school had a 7.00, and the two remaining schools' had 8.00. The median values ranged from 6.00 to 8.50. School #1 had the lowest mean score at 6.55; the highest mean score was for school #6 at 8.45. The standard deviations ranged from 1.33 (school #2) to 1.93 (school #5). The coefficient of variation for school #2 was lowest at .163, while the highest value was school #5 at .276. The standard errors ranged from .191 for school #4 to .483 for school #5. The skewness values were split with three being positive and three being negative, with the smallest at .257 and the largest at -.785. Dimension #2 consisted of three items for a maximum score of 15 points, which all six schools obtained. The minimum scores

Table 5

Descriptive Statistics for the First Four Dimension Scores
by the Six Middle/Junior High Schools

School Number	Number of Teachers	Minimum Score Obtained	Maximum Score Obtained	Mode	Median	Mean	Standard Deviation	Coeff. of Variation	Standard Error	Skew-ness
1. Principal's Facilitative Leadership										
1	40	3.00	10.00	6.00	6.00	6.55	1.45	.221	.229	.428
2	17	5.00	10.00	8.00	8.00	8.18	1.33	.163	.324	-.721
3	32	4.00	10.00	6.00	6.50	6.84	1.73	.253	.305	.257
4	50	4.00	10.00	6.00	6.00	6.66	1.35	.203	.191	.502
5	16	3.00	10.00	7.00	7.00	7.00	1.93	.276	.483	-.760
6	20	5.00	10.00	8.00	8.50	8.45	1.43	.169	.324	-.785
2. Shared Visions for Improvement										
1	40	6.00	15.00	9.00	10.00	10.43	2.17	.208	.343	.400
2	17	9.00	15.00	14.00	14.00	13.24	1.64	.124	.398	-1.101
3	32	7.00	15.00	13.00	12.00	12.00	1.87	.156	.330	-.667
4	52	5.00	15.00	12.00	11.00	10.90	2.42	.222	.336	-.367
5	16	8.00	15.00	8.00	11.50	11.25	2.70	.240	.674	.047
6	21	8.00	15.00	13.00	13.00	12.81	2.02	.158	.440	-.847
3. Collective Creativity and Learning										
1	39	7.00	25.00	16.00	16.00	16.82	4.03	.240	.645	.044
2	17	18.00	25.00	21.00	21.00	20.88	2.18	.104	.528	.706
3	33	15.00	24.00	16.00	20.00	19.49	2.51	.129	.438	-.133
4	51	8.00	25.00	20.00	19.00	18.41	3.33	.181	.466	-.461
5	16	11.00	23.00	18.00	19.00	18.06	3.40	.188	.849	-.797
6	20	12.00	25.00	23.00	21.00	20.20	3.62	.179	.810	-.711
4. Classroom Observations and Feedback										
1	40	2.00	10.00	4.00	4.00	4.50	1.83	.407	.289	.564
2	14	2.00	7.00	4.00	5.00	4.79	1.58	.330	.422	-.413
3	30	2.00	8.00	6.00	5.50	5.00	1.91	.382	.349	-.095
4	47	2.00	10.00	4.00	5.00	4.98	1.80	.361	.263	.758
5	16	2.00	8.00	2.00	4.00	4.50	2.00	.444	.500	.171
6	19	2.00	8.00	2.00	4.00	4.37	2.03	.465	.466	.235

Table 6

Descriptive Statistics for the Fifth Dimension Score and the Total Instrument Score by the Six Middle/Junior High Schools

School Number	Number of Teachers	Minimum Score Obtained	Maximum Score Obtained	Mode	Median	Mean	Standard Deviation	Coeff. of Variation	Standard Error	Skew-ness
5. School Conditions and Capacities										
1	39	7.00	25.00	13.00	15.00	15.10	4.12	.273	.660	.351
2	17	18.00	25.00	21.00	22.00	21.88	2.37	.108	.574	-.289
3	32	12.00	25.00	21.00	19.00	18.88	3.02	.160	.535	-.194
4	51	8.00	25.00	17.00	17.00	17.04	3.69	.217	.517	-.290
5	15	12.00	21.00	18.00	17.00	16.47	2.64	.160	.682	-.319
6	21	13.00	25.00	23.00	20.00	20.14	3.15	.156	.688	-.410
Total Instrument Score										
1	38	31.00	85.00	47.00	53.00	53.32	11.97	.224	1.942	.435
2	14	63.00	77.00	63.00	69.00	69.57	4.36	.063	1.166	.287
3	28	42.00	79.00	63.00	63.00	62.14	8.94	.144	1.689	-.583
4	46	33.00	85.00	61.00	58.00	57.57	10.59	.184	1.561	.155
5	15	37.00	70.00	70.00	59.00	56.87	10.88	.191	2.808	-.681
6	17	48.00	77.00	75.00	70.00	66.47	9.75	.147	2.366	-.853

ranged from 5.00 to 9.00. The modes ranged from 8.00 to 14.00, while the medians ranged from 10.00 to 14.00. The highest mean score and smallest standard deviation were for school #2 at 13.24 and 1.64, respectively. School #1 had the lowest mean score at 10.43, followed closely by school #4 at 10.90. The largest standard deviation (2.70) was for school #5, which contributed to that school having the largest coefficient of variation value at .240. School #5 also had the highest standard error at .674—considerably higher than all others. Regarding the skewness values, four were negative with -1.101 for school #2 being the largest value. Of the two positive skewness values, .047 for school #5 was the closest of all six to being normal.

Table 5 also displays the descriptive data for dimension #3, a 25-point maximum score, and dimension #4, a 10-point maximum score. For dimension #3, the minimum scores for the six schools ranged from 7.00 to 18.00 and the maximum scores were a much more narrow range of 23.00 to 25.00, with four of the six having maximums at 25.00. The mode scores ranged from the low of 16.00 for two schools up to 23.00 for school #6. The median scores ranged from a low of 16.00 for school #1 up to 21.00 for schools #2 and #6. Similarly, the means ranged from 16.82 for school #1 to 20.88 for school #2. School #2 had both the lowest standard deviation at 2.18 and coefficient of variation at .104. School #1 had the largest standard deviation at 4.03 and, also, the largest coefficient of variation at .240. The standard errors varied from the low of .438 (school #3) to the high of .849 (school #5). Four of the six skewness values were negative and two were positive; none was over 1.0. The two positive skewness values were very different with school #2 at .706 and school #1 at .044, close to normal distribution skewness. For dimension #4, all six schools had the lowest possible minimum score of 2.00, while the maximum scores varied from 7.00 to 10.00. The modes also varied considerably with two (schools #5 and #6) at 2.00, three (schools #1, #2, and #4) at 4.00 and the remaining school (#3) at 6.00. The medians varied much less from 4.00 to 5.50. Interestingly, all six mean scores were at the mid-point value of 5.00 or less, with school #3 at the mid-point. The standard deviations ranged from 1.58 to 2.03, which resulted in all six coefficient of variation values being higher than any others in Table 5. Specifically, the coefficient of variation values ranged from .330 (school #2) up to .465 (school #6). The standard errors ranged from .263 to .500. In contrast to the prior dimension, four of six skewness values were positive and one of the two negative values (-.095 for school #3) approached normalcy.

Table 6 displays the descriptive data for dimension #5, a 25-point maximum score. The minimum scores ranged widely from 7.00 to 18.00, while the maximum scores ranged only from 21.00 to 25.00. The mode scores also ranged widely from 13.00 for school #1 to 23.00 for school #6. The median scores ranged rather widely from 15.00 to 22.00. Similarly, the mean scores ranged rather widely from 15.10 for school #1 to 21.88 for school #2. School #2 also had the smallest standard deviation at 2.37 and the smallest coefficient of variation at .108. School #1 had the largest standard deviation at 4.12 and the largest coefficient of variation at .273. The standard errors ranged from .517 to .688. The skewness values ranged from .194 to .410, all were negative values except school #1.

Finally, for the six middle/junior high schools, Table 6 displays the descriptive statistics for the 17 to 85-point total instrument score. The minimum scores for the six schools ranged widely from 31.00 to 63.00 with three in the 30s, two in the 40s, and one in the 60s. The range value for

the maximum scores was 15 points, from 70.00 for school #5 to 85.00 for schools #1 and #4. The modes ranged widely from 47.00 to 75.00, while the medians varied less from 53.00 (school #1) to 70.00 (school #6). The mean scores and the standard deviations revealed large differences across the six schools on the total score. As an example, school #2 was the highest scoring school of the group with a mean of 69.57, standard deviation of 4.36, and a coefficient of variation of .063. School #6 was second highest with a mean of 66.47, a standard deviation of 9.75, and a coefficient of variation of .147. In sharp contrast to those two, school #1 was the lowest scoring of the group on the total instrument with a mean of 53.32, a standard deviation of 11.97 (nearly triple that of school #2), and a coefficient of variation of .224—highest of the six and the only one over .200. The standard errors ranged from a low of 1.166 for school #2 to 2.808 for school #5. In terms of the skewness values for the six schools, three were positive and three were negative. The skewness values ranged from .155 to .853.

Tables 7 and 8 display the descriptive statistics for the five dimensions and the total score for the nine high schools in the field test. The data for the first three dimensions are presented in Table 7 and the data for the last two dimensions and the total score are presented in Table 8. The number of teachers in the nine high schools ranged from a low of 11 to a high of 60. Two schools (#6 and #9) had faculties under 20 in number, while four schools (#1, #2, #3, and #4) had faculties in the 50s and 60s.

Dimension #1, Principal's Facilitative Leadership, consisted of two items for possible scores of 2.00 to 10.00. Table 7 shows that the minimum scores for the nine schools ranged from 2.00 to 7.00 points for school #6. The maximum scores were from 8.00 for school #2 to 10.00 for all other schools. The mode scores ranged from 4.00 for two schools up to 8.00 for school #6. Regarding the median scores, for the first five high schools, it was 6.00; for the remaining four high schools, it was 8.00. School #6 had the highest mean, smallest standard deviation, and smallest coefficient of variation at 8.62, 1.21, and .140, respectively. School #2 had the lowest mean score at 5.19. Two other schools had means under 6.00, two had means under 7.00, and three had means under 8.00. The standard deviations ranged from the aforementioned 1.21 (school #6) up to 1.98 (school #4). All of the coefficient of variation values were in the .200s and .300s except for school #6 at .140. The standard errors ranged from .198 for school #1 to .400 for school #9. Four of the nine skewness values were negative and five were positive. Interestingly, two negative (schools #8 and #9) and two positive (schools #6 and #7) were less than .100, indicating close to normal skewness for the distribution of scores.

Dimension #2, Shared Visions for Improvement for Improvement, has three items for a 15-point maximum score. The minimum scores ranged widely from a low of 3.00 for schools #2 and #4 to a high of 11.00 for school #6. The maximum scores were much closer, ranging from 13.00 for school #2 to 15.00 for six of the nine high schools. The modes ranged from 9.00 for four schools to 15.00 for school #8. The median values ranged from 8.00 for school #2 to 13.00 for school #8. Similar to the first dimension above, school #6 had the highest mean, the smallest standard deviation, and the smallest coefficient of variation at 12.87, 1.30, and .101, respectively. Schools #8 and #9 were close to school #6 with means of 12.57 and 12.47, respectively, but also with larger standard deviations. The lowest mean score of 8.10 was for school #2 with a standard deviation of 2.12,

Table 7

Descriptive Statistics for the First Three Dimension Scores
by the Nine High Schools

School Number	Number of Teachers	Minimum Score Obtained	Maximum Score Obtained	Mode	Median	Mean	Standard Deviation	Coeff. of Variation	Standard Error	Skew-ness
1. Principal's Facilitative Leadership										
1	51	2.00	10.00	6.00	6.00	6.35	1.41	.222	.198	-.176
2	57	2.00	8.00	6.00	6.00	5.19	1.61	.310	.213	-.246
3	60	2.00	10.00	4.00	6.00	5.83	1.92	.329	.247	.231
4	58	2.00	10.00	5.00	6.00	5.95	1.98	.333	.260	.469
5	46	4.00	10.00	4.00	6.00	6.15	1.62	.263	.239	.137
6	13	7.00	10.00	8.00	8.00	8.62	1.21	.140	.311	.079
7	25	5.00	10.00	6.00	8.00	7.56	1.73	.229	.347	.073
8	34	5.00	10.00	6.00	8.00	7.74	1.66	.214	.284	-.017
9	19	4.00	10.00	6.00	8.00	7.42	1.74	.235	.400	-.099
2. Shared Visions for Improvement										
1	51	6.00	14.00	9.00	10.00	9.78	1.69	.173	.237	-.113
2	58	3.00	13.00	9.00	8.00	8.10	2.12	.262	.278	-.003
3	60	5.00	15.00	9.00	9.00	9.85	2.28	.231	.294	.307
4	58	3.00	15.00	12.00	10.00	10.14	2.41	.238	.316	-.589
5	43	5.00	15.00	9.00	11.00	10.35	2.25	.217	.342	-.102
6	15	11.00	15.00	12.00	12.00	12.87	1.30	.101	.336	.729
7	25	5.00	14.00	13.00	11.00	10.84	2.59	.239	.519	-.900
8	35	7.00	15.00	15.00	13.00	12.57	2.23	.177	.376	-.694
9	19	9.00	15.00	12.00	12.00	12.47	1.74	.140	.400	.014
3. Collective Creativity and Learning										
1	51	7.00	22.00	17.00	16.00	15.53	3.41	.220	.477	-.314
2	56	5.00	22.00	12.00	14.00	14.05	3.46	.246	.463	.057
3	58	7.00	25.00	17.00	17.00	15.88	3.97	.250	.522	-.328
4	56	6.00	24.00	17.00	17.00	16.00	3.56	.223	.475	-.269
5	45	7.00	23.00	19.00	18.00	16.82	3.38	.201	.504	-.727
6	15	16.00	25.00	19.00	20.00	20.60	2.80	.136	.722	.453
7	25	13.00	24.00	20.00	19.00	18.56	3.07	.165	.614	.162
8	34	14.00	25.00	20.00	20.00	19.74	2.91	.147	.498	-.240
9	18	14.00	25.00	20.00	19.50	18.89	2.91	.147	.685	.220

Table 8

Descriptive Statistics for the Last Two Dimension Scores and the Total Instrument Score by the Nine High Schools

School Number	Number of Teachers	Minimum Score Obtained	Maximum Score Obtained	Mode	Median	Mean	Standard Deviation	Coeff. of Variation	Standard Error	Skew-ness
4. Classroom Observations and Feedback										
1	49	2.00	9.00	2.00	6.00	4.94	2.16	.437	.309	-.149
2	55	2.00	9.00	2.00	4.00	4.35	2.11	.485	.285	.433
3	55	2.00	10.00	4.00	5.00	5.16	2.26	.438	.305	.299
4	59	2.00	9.00	2.00	5.00	4.85	2.24	.462	.292	.064
5	43	2.00	10.00	4.00	6.00	5.26	1.92	.365	.292	.021
6	12	4.00	8.00	6.00	6.00	5.83	1.38	.237	.386	-.188
7	24	2.00	10.00	5.00	4.50	4.79	2.21	.461	.450	.844
8	33	2.00	10.00	4.00	7.00	5.97	2.19	.367	.381	-.341
9	19	4.00	8.00	8.00	7.00	6.58	1.39	.211	.318	-.531
5. School Conditions and Capacities										
1	48	10.00	24.00	13.00	15.00	15.38	3.02	.196	.436	.430
2	54	5.00	22.00	16.00	14.00	13.50	3.65	.270	.496	.119
3	55	8.00	24.00	17.00	17.00	16.60	3.90	.235	.526	-.268
4	52	5.00	25.00	15.00	15.00	16.00	4.84	.303	.672	-.058
5	43	7.00	25.00	20.00	17.00	17.26	3.69	.214	.563	-.273
6	15	13.00	25.00	25.00	20.00	20.33	3.29	.162	.849	-.341
7	24	8.00	25.00	20.00	16.00	15.63	4.54	.290	.926	.029
8	33	7.00	25.00	17.00	19.00	18.52	4.09	.221	.713	-.984
9	19	13.00	25.00	16.00	19.00	18.58	3.45	.186	.792	.336
Total Instrument Score										
1	42	35.00	74.00	50.00	51.50	51.95	8.42	.162	1.299	.119
2	48	17.00	70.00	51.00	45.00	45.21	10.59	.234	1.528	-.052
3	49	26.00	76.00	51.00	52.00	52.96	11.76	.220	1.680	-.218
4	49	24.00	80.00	55.00	54.00	53.69	13.08	.244	1.869	.092
5	39	25.00	77.00	56.00	56.00	55.69	11.46	.206	1.835	-.385
6	11	62.00	82.00	82.00	67.00	70.36	8.41	.120	2.534	.563
7	24	39.00	83.00	41.00	56.50	57.17	12.41	.217	2.532	.166
8	30	36.00	85.00	68.00	68.00	65.07	10.84	.167	1.980	-.850
9	18	49.00	83.00	57.00	63.00	64.00	8.98	.140	2.116	.397

yielding the largest coefficient of variation value at .262. The standard errors for the schools were from a low of .237 for school #1 to a high of .519 for school #7. Six of the skewness values were negative and the remaining three were positive. At -.003, the skewness for school #2 was almost exactly normal and, at .014, the skewness for school #9 was close to normal.

Dimension #3, Collective Creativity and Learning, had five items with a 25-point maximum score. The minimum scores for the nine schools ranged from the lowest possible score of 5.00 for school #2 to 16.00 for school #6. The maximum scores were in a more narrow range of 22.00 (two schools) to 25.00 (four schools). School #2 had the lowest mode at 12.00, while schools #7, #8, and #9 had the highest at 20.00. The median scores varied from 14.00 for school #2 to 20.00 for schools #6 and #8. Consistent with the pattern established in the two prior dimensions, school #6 had the highest mean (20.60) and lowest standard deviation and coefficient of variation, while school #2 had the lowest mean score (14.05) and close to the largest standard deviation and coefficient of variation values. The standard deviations ranged from school #6's low of 2.80 to school #3's high of 3.97. The coefficient of variation values were from .136 for school #6 to .250 for school #3. The standard errors ranged from .463 (school #2) to .722 (school #6). With respect to the skewness values, five were negative and four were positive. Only school #2 (.057) was close to normal in skewness of scores.

Table 8 displays the descriptive data for dimension #4, Classroom Observations and Feedback, which has a two-item, 10-point maximum score. Seven of the nine schools produced the lowest possible minimum score of 2.00, while the maximum scores ranged from 8.00 (two schools) to 10.00 (four schools). The mode scores varied widely from 2.00 (three schools) to 8.00 (school #9). The median scores varied from 4.00 (school #2) to 7.00 (schools #8 and #9). The range value for the mean scores was 2.23, from a low of 4.35 for school #2 to a high of 6.58 for school #9. School #9 also had the second smallest standard deviation at 1.39 (.01 larger than the smallest) and the smallest coefficient of variation at .211. Consistent with the prior dimensions, school #2 had the lowest mean score (4.35) and the largest coefficient of variation value at .485. School #9 had the highest mean at 6.58, while school #6 had the largest standard deviation of 2.26. The standard errors ranged from a low of .285 for school #2 to a high of .450 for school #7. Five of the skewness values were positive and the remainder were negative. Two of the positive skewness values indicated the distribution of scores for schools #4 (.064) and #5 (.021) were close to normal.

Dimension #5, School Conditions and Capacities, has five items, and a 25-point maximum score. One school (#4) obtained the lowest possible minimum score of 5.00, while the last six schools obtained the maximum possible score of 25.00. The mode scores ranged from a low of 13.00 (school #1) to a high of 25.00 (school #6). The median scores ranged from 14 (school #2) to 20.00 (school #6). School #6 also had the highest mean score at 20.33 and the smallest coefficient of variation at .162. Consistent with prior dimensions, school #2 had the lowest mean score of 13.50, producing a large range value of 6.83 on the 20-point possible range score. Six of the standard deviations were under 4.00, and the largest was 4.84 for school #4. The coefficient of variations differed from a low of .162 to a high of .303 (school #4). The standard errors varied considerably from .436 (school #1) to .926 (school #7). Five of the skewness values were negative and one of those (school #4 at -.058) was close to normal. Of the four positive skewness values, school #7 at .029 was close to normal.

Finally, Table 8 displays the descriptive data for the 17-item, 85-point maximum total score. One school (#2) produced the lowest possible minimum score of 17.00, while the minimum score for school #6 was 62.00. The maximum scores for the nine schools ranged from the low of 70.00 for school #2 to the high of 85.00 for school #8. In fact, five of the nine maximum scores were in the 80s. The mode scores ranged from 50.00 (school #1) to 82.00 (school #6). The median scores varied from the low of 45.00 for school #2 to the high of 68.00 for school #8. Consistent with the pattern established for most of the dimension scores above, school #6 had the highest mean score, the smallest standard deviation, and the smallest coefficient of variation at 70.36, 8.41, and .120, respectively. Also consistent with all five dimensions, school #2 had the lowest mean score at 45.21, which is 6.74 points less than the next lowest mean score. School #4 had the largest standard deviation (13.08), while school #4 had the largest coefficient of variation (.244). The standard errors ranged from 1.299 (school #1) to 2.534 (school #6). Five of the skewness values were positive and four were negative and ranged from -.850 (school #8) to .563 (school #6). School #2 (-.052) and school #4 (.092) were close to normal in terms of skewness.

Usability. The usability of the reformatted Hord instrument was assessed two different ways. First, data in Tables 3 to 8, above, show that the instrument does differentiate among the schools on its five dimensions and total score. This usability function of differentiating schools on its six scores also holds true when the schools were grouped by their level (elementary, middle/junior high, and high school). Second, the usability of the instrument was determined on the basis of its "workability," how well it "worked" in the field test. Here we were very interested in the practical aspects of its administration and completion by the teachers in the 21 schools. For example, we found at data entry time that only one teacher skipped the second page and just two teachers failed to complete the third, and last, page. In terms of the individual items in the instrument, Table 1 shows most items being completed by the vast majority of the teachers. The item skipped the most by the full group was #4b, which was skipped by 41 of the 690 teachers. Data entry clerks noted less than a dozen instruments in which the respondents did not follow directions to mark, circle, or check one of the five numbers directly, but marked somewhere between two numbers on the scale. (In these few cases, data entry clerks were instructed to measure which number the mark was closest to and enter that number.)

Last, we do know that the unique response scale, with three complete sentences to define the end points and the middle of the scale, required a little more time for reading by the respondents and, thus, a little more administration time than other instruments with different response options. One advantage of the instrument's response format was that the evaluators noticed much less writing of notes, comments, or questions for the items than is usual for an instrument completed by a group of teachers. Less than two dozen instruments had teacher-supplied notes, comments, or questions for any item. This is judged to be another indication of the instrument's usability.

Reliability

Internal Consistency Reliability

Table 9 displays the correlations among the instrument's 17 items by the full group of teachers in the field test. There was complete data for all items for 595 teachers. All of the correlations in the table are significant at the .0001 level. Also, all of the correlations among the 17 individual items were positive. These positive and significant correlations ranged from the low of .3060 (item 4b with item 2a) to the high of .7862 (item 1b with item 1a). In Table 9, there are 22 correlations in the .30s, 61 in the .40s, 40 in the 50s, 9 in the 60s, and just 4 in the .70s. The trend was for all of the highest correlations (.60s and .70s) to be items within the same dimension, except the .6148 correlation of item 5a with item 3b. Another trend was that the majority of the lowest correlations included items 4a and 4b, the pair of items comprising the Classroom Observations and Feedback dimension. The last trend noted was that the vast majority of the intercorrelations of the 17 items were moderate, being in the .40s and .50s.

Table 10 displays the Cronbach Alpha reliabilities and associated statistics for the five individual dimensions in the instrument. The internal consistency reliability (Alpha) for dimension #1, Principal's Facilitative Leadership, was .8703. With two items, the mean was under 3.50 if either one was omitted from the dimension. Item 1b had the larger variance, but both items had the same high correlation (.7765) with the total instrument. The Alpha reliability for dimension #2, Shared Visions for Improvement, was .8304 for its three items. Item 2a had the largest variance of the trio, the lowest correlation with the total instrument score, and contributed less than the other two items to the Alpha coefficient. The third dimension, Collective Creativity and Learning, had an Alpha reliability coefficient of .8601 for its five items. Item 3d had the most variance of the five, the highest correlation with the total score (.7015), and affected the Alpha figure the most if it were to be deleted from the dimension. The Alpha reliability for dimension #4, Classroom Observations and Feedback, was .8434 for its pair of items. Item 4a produced more variance than item 4b, and both items had a correlation of .7418 with the total instrument. The fifth dimension, School Conditions and Capacities, had an Alpha reliability of .8489 for its five items. Of those five, item 5d yielded the most variance (12.3798), had the lowest correlation with the total score, and affected the Alpha figure the least if it were to be deleted from the dimension. In contrast, Item 5b had the highest correlation (.6896) with the total score and affected the Alpha coefficient the most if it were to be deleted. In summary, all of the Alpha reliability coefficients for the five dimensions were in the mid .80s, ranging from .8304 up to .8703.

Table 11 displays the Cronbach Alpha reliability coefficients for the total instrument together with the item statistics for the 17 items. The internal consistency reliability (Alpha) for the total instrument was .9389. Interestingly, the data in the last column show that this Alpha figure would increase if only one item (4b) was deleted from the set—nearly all the items contributed equally to the rather high Alpha coefficient for the total instrument. The two items with the lowest total mean scores if they were deleted were 2b and 3d, while the pair of items with the highest total mean scores if deleted were 4a and 4b. Item 5a contributed the most to the total score variance, followed next

Table 9

Correlations Among the 17 Individual Items by the Full Group

Item No.	1a	1b	2a	2b	2c	3a	3b	3c	3d	3e	4a	4b	5a	5b	5c	5d	5e
1a	1.0000																
1b	.7862	1.0000															
2a	.5807	.5708	1.0000														
2b	.5453	.5272	.6185	1.0000													
2c	.5568	.5283	.6135	.6835	1.0000												
3a	.4780	.4898	.4803	.4782	.5444	1.0000											
3b	.5450	.5388	.5131	.5283	.5555	.6328	1.0000										
3c	.4717	.4668	.5177	.5246	.5157	.4819	.5781	1.0000									
3d	.4239	.4147	.4753	.5272	.5440	.4667	.5395	.5829	1.0000								
3e	.4218	.4404	.5152	.5140	.5271	.4842	.5152	.6066	.7049	1.0000							
4a	.3523	.3192	.3333	.3433	.3791	.3886	.4095	.3840	.3440	.4051	1.0000						
4b	.3222	.3327	.3060	.3063	.3619	.3189	.3341	.3489	.3303	.4233	.7344	1.0000					
5a	.4494	.4416	.4390	.5233	.5061	.5150	.6148	.4974	.4653	.4986	.3817	.3752	1.0000				
5b	.4715	.4501	.4610	.4902	.4898	.4810	.5330	.4982	.4592	.4906	.4189	.4132	.6434	1.0000			
5c	.5589	.5060	.5103	.5378	.5458	.4691	.5555	.5447	.5447	.5013	.4408	.4255	.6083	.6052	1.0000		
5d	.4059	.3944	.4265	.4639	.4755	.4256	.3732	.4002	.4145	.4433	.4344	.4114	.3794	.4454	.4741	1.0000	
5e	.4129	.4103	.4479	.4868	.4961	.4779	.4588	.4654	.4690	.5133	.4340	.4044	.4613	.5233	.5406	.7793	1.0000

Note: The number of respondents in each cell is 595 and all correlations are significant at the .0001 level.

Table 10

Cronbach's Coefficient Alpha Reliability with Item Correlations for the Five Individual Dimensions by the Full Group

Item Number	Dimension Mean if Item Deleted	Dimension Variance if Item Deleted	Correlation with Total Instrument	Alpha if Item Deleted
1. Principal's Facilitative Leadership, Alpha = .8703				
1a	3.4121[a]	0.8699	.7765	N/A
1b	3.3117	1.0965	.7765	N/A
2. Shared Visions for Improvement, Alpha = .8304				
2a	7.4742[b]	3.2261	.6464	.8072
2b	7.1178	3.0805	.7124	.7420
2c	7.3520	3.0898	.7094	.7452
3. Collective Creativity and Learning, Alpha = .8601				
3a	14.4086[c]	9.9474	.6200	.8478
3b	14.3997	9.9492	.6936	.8268
3c	14.1857	10.1931	.6886	.8282
3d	13.9569	10.2794	.7015	.8256
3e	14.1100	10.2736	.6959	.8268
4. Classroom Observations and Feedback, Alpha = .8434				
4a	2.7473[d]	1.4268	.7418	N/A
4b	2.3652	0.9853	.7418	N/A
5. School Conditions and Capacities, Alpha = .8489				
5a	13.7176[e]	10.9828	.6376	.8264
5b	13.9588	11.2689	.6896	.8093
5c	13.7634	11.7272	.6843	.8115
5d	13.8611	12.3798	.6029	.8320
5e	13.7008	11.6198	.6900	.8098

[a]N=677.
[b]N=679.
[c]N=673.
[d]N=649.
[e]N=655.

Table 11

Cronbach's Coefficient Alpha Reliability with Item Correlations for the Total Instrument by the Full Group

Item Number	Total Score Mean if Item Deleted	Total Score Variance if Item Deleted	Correlation with Total Instrument	Alpha if Item Deleted
Total Instrument Score, Alpha = .9389				
1a	54.6605	134.3324	.6770	.9351
1b	54.5580	136.4693	.6652	.9354
2a	54.4487	135.3320	.6790	.9351
2b	54.1210	134.7833	.7066	.9345
2c	54.3412	134.6292	.7283	.9341
3a	54.6134	134.4867	.6639	.9345
3b	54.5899	133.6161	.7215	.9341
3c	54.3782	135.9965	.6872	.9350
3d	54.1429	136.9812	.6596	.9355
3e	54.3092	136.0557	.6983	.9348
4a	55.6000	137.6949	.5699	.9374
4b	55.2067	135.7703	.5306	.9390
5a	54.4336	131.9059	.6833	.9351
5b	54.6588	133.4575	.6915	.9348
5c	54.4773	134.0378	.7274	.9340
5d	54.5933	136.6390	.6219	.9363
5e	54.4370	134.8222	.6794	.9350

Note: Total number of respondents with full data was 595.

by item 5b. In contrast, item 4a contributed the least to the total score variance, followed by item 3d. The two dimension #4 items correlated the lowest with the total instrument at .5306 for 4b and .5699 for 4a. Item 2c correlated the highest with the total score at .7283, while the next highest was 5c at .7274. The last column shows that the impact of deleting any one of the 17 items is very small on the overall Alpha coefficient, ranging from .9340 to .9390. If item 4b was deleted, the Alpha coefficient would go up .0001 of a point to .9390. However, if item 5c was deleted, the Alpha would drop slightly to .9340, and to .9341 if either item 2c or 3b was deleted.

Table 12 displays the Cronbach Alpha internal consistency reliabilities for the five dimensions and the total instrument for the 21 schools. The level of the school and the number of teachers who completed all 17 items are presented in the second and third columns. The Alpha reliabilities for the two-item dimension #1 ranged from .68 (a middle/junior high school) to .91 (a high school) with 14 Alphas in the .80s and 4 in the .70s. The SPSS program did not compute an Alpha reliability coefficient for school #14 (an elementary school) because all teachers responded with the highest possible option, thus producing no variance across the two items. The Alpha reliabilities for the three-item dimension #2 ranged from .52 (a high school) to .91 (an elementary school). There was one other Alpha coefficient in the .50s and two in the .60s; the remainder were in the .70s and .80s. Dimension #3 consisted of five items and its Alpha coefficients ranged from .56 (an elementary school) to .91 (a middle/junior high school). There was another Alpha coefficient in the .50s and one in the .60s; the remainder were in the .70s and 80s.

Regarding the two-item dimension #4, the Alphas were very similar to the previous dimension with two in the 50s (an elementary school and a high school), one in the .60s, most in the .70s and .80s, and one in the .90s (a middle/junior high school). Dimension #5, with five items, had one Alpha at .59 (a middle/junior high school), one at .66 (an elementary school), seven in the .70s, and the remainder n the .80s, including four tied at .88 for the highest Alpha coefficient. Finally, the last column reveals a range of total instrument Alpha reliability coefficients from .62 (a middle/junior high school) up to .95 (for two high schools). All of the schools (except the one with a .62 coefficient) were split with 7 Alphas in the .80s and 13 in the 90s. To summarize the data in Table 12, of the 126 Alpha reliabilities displayed, 7 were in the .50s, 7 were in the .60s, 33 were in the .70s, 61 were in the .80s, 17 were in the .90s, and 1 was not computed.

Table 13 displays the correlations among the five dimension scores and the total instrument score by the full group of 595 teachers with complete data. All of the correlations are positive and, further, all of them are significant at the .0001 level. The correlations in Table 13 ranged from the low of .3764 (dimension #4 with dimension #1) to the high of .9096 (the total score with dimension #5). Overall, the correlation between the dimensions and the total score were higher (.7561, .8534, .9045, .6541, and .9096 in order) than among the dimensions themselves. Dimension #4, Classroom Observations and Feedback, had the four lowest correlations in Table 12, with dimensions #1, #2, #3, and #5. However, at .6541, the correlation of dimension #4 to the total score was higher than with the other dimensions.

Table 12

Cronbach Alpha Internal Consistency Reliabilities for the Five Dimensions and the Total Instrument for the 21 Schools

School No.	School Level[a]	Number of Teachers[b]	Dimension Names and Number of Items and the Total Instrument					
			1. Principal's Facilitative Leadership (2 Items)	2. Shared Visions for Improvement (3 Items)	3. Collective Creativity and Learning (5 Items)	4. Classroom Observations and Feedback (2 Items)	5. School Conditions and Capacities (5 Items)	Total Instrument (17 Items)
1	Mid/Jr #1[c]	38	.68	.77	.86	.80	.82	.94
2	H.S. #1	42	.86	.52	.80	.89	.70	.86
3	H.S. #2	48	.83	.78	.77	.83	.81	.91
4	H.S. #3	49	.87	.70	.85	.87	.85	.93
5	H.S. #4	49	.81	.78	.78	.88	.88	.94
6	Mid/Jr. #2	14	.83	.76	.68	.87	.76	.62
7	H.S. #5	39	.84	.78	.86	.87	.84	.95
8	H.S. #6	11	.73	.68	.86	.54	.88	.95
9	H.S. #7	24	.83	.77	.75	.82	.88	.94
10	H.S. #8	30	.86	.87	.76	.84	.87	.93
11	Mid/Jr. #3	28	.80	.73	.72	.84	.79	.90
12	Elem. #1	27	.73	.57	.80	.52	.70	.84
13	Elem. #2	38	.74	.66	.83	.73	.78	.89
14	Elem. #3	15	N/A[d]	.86	.78	.89	.66	.87
15	Mid/Jr. #4	46	.84	.84	.86	.83	.85	.94
16	Mid/Jr. #5	14	.89	.82	.83	.77	.59	.93
17	H.S. #9	18	.91	.78	.80	.74	.82	.91
18	Mid/Jr. #6	17	.82	.86	.91	.94	.88	.93
19	Elem. #4	27	.82	.71	.77	.77	.73	.85
20	Elem. #5	8	.70	.77	.59	.66	.82	.86
21	Elem. #6	13	.88	.91	.56	.88	.71	.87

[a]Labels: Elem. = Elementary school, Mid/Jr. = Middle/Junior high school, H.S. = High school.

[b]This is the number of teachers who completed all 17 items in the instrument, which usually is less than the number of teachers completing the items for the five dimensions.

[c]The number after the level label corresponds to the order of the schools in the previous tables. For example, Mid/Jr. #1 is the same school as school number 1 in Tables 5 and 6.

[d]N/A means there was no variance in the two items as all teachers responded with the same, highest possible response.

Table 13

Correlations Among the Five Dimension Scores and the Total Instrument Score by the Full Group

Dimension or Total Score Name	Dimension and Total Score Names					
	1. Principal's Facilitative Leadership	2. Shared Visions for Improvement	3. Collective Creativity and Learning	4. Classroom Observations and Feedback	5. School Conditions and Capacities	Total Instrument Score
1. Principal's Facilitative Leadership	1.0000					
2. Shared Visions for Improvement	.6705	1.0000				
3. Collective Creativity and Learning	.6201	.7384	1.0000			
4. Classroom Observations and Feedback	.3764	.4150	.4895	1.0000		
5. School Conditions and Capacities	.5976	.7001	.7521	.5537	1.0000	
Total Instrument Score	.7561	.8534	.9045	.6541	.9096	1.0000

Note: The number of respondents in each cell is 595 and all correlations are significant at the .0001 level.

Stability Reliability

Table 14 displays the stability (test-retest) descriptive statistics and reliability coefficients for the five dimensions and the total score by the three high school faculties in this subsample of the field test. The number of teachers participating in the first administration of the instrument ranged from 112 to 122 on the individual dimensions and was 101 for the total instrument. The number of teachers participating in the second administration (retest) ranged from 61 to 64 on the dimensions and was 58 on the total instrument. However, the maximum number of teachers that could be matched by individually-supplied identification numbers was 30 on dimension #2 and was just 23 on the total instrument score. Data in the mean score column show that three of the five dimension means increased slightly at less than .40, while two declined slightly at less than .70. On the 85-point maximum total instrument score, the score declined .82 to 50.38 at the second administration. For the six standard deviations, four increased in size from the first to second administration and the other two (dimensions #1 and #4) decreased. The stability coefficients for dimensions #1, #2, #3, and #5 were in the .40s, while for dimension #4 it was .5657. The stability coefficient for the total instrument score was .6147.

Validity

Content Validity

The content validity of the instrument was determined at three stages. In the first stage, the content of the five dimensions, the 17 descriptors, and the 51 indicators was established by the author (Hord) when she constructed them. The content for the dimensions, descriptors, and indicators was derived from her extensive review of the educational and business/corporate literature, plus her field research with selected schools in the southwest United States that functioned as professional learning communities. As described earlier, Hord was influenced greatly by her long-term qualitative research with one particular elementary school that continued operating as a professional learning community through the change of four principals in 15 years.

The second stage of the content validity assessment was completed by three AEL staff as part of the reformatting of the instrument for use in its pilot test in an AEL project. This stage of the content review was described in detail in the Methodology section and will not be repeated here. Basically, every descriptor and indicator was reviewed independently by the three AEL coauthors and the instrument was reformatted after consensus on wording was reached. AEL's review attempted to hold true to the original intent and meaning of Hord's version, yet apply some additional clarification and consistency, especially across the trio of indicators under each response scale.

The third stage of content review was completed when AEL staff sent their reformatted instrument to the original developer for another check on its content. The purpose of this review stage was for Hord to assess the minor word changes made by AEL staff and also to confirm that the reformatting was consistent with her original intentions for the instrument. Hord's review confirmed that the changes were satisfactory to her and that the reformatting did continue the intentions she had

Table 14

Stability Descriptive Statistics and Reliability Coefficients for the Five Dimensions and Total Score by the Three High School Faculties

Dimension Number and Names and Total Instrument	Instrument Administration	Number	Mean	Standard Deviation	Number Test-Retest Matches	Coefficient of Stability
1. Principal's Facilitative Leadership	First Admin. Second Admin.	120 63	5.83 5.94	1.83 1.73	29	.4344
2. Shared Visions for Improvement	First Admin. Second Admin.	122 64	9.42 9.80	2.34 2.66	30	.4515
3. Collective Creativity and Learning	First Admin. Second Admin.	118 61	15.36 14.75	3.30 3.57	27	.4154
4. Classroom Observations and Feedback	First Admin. Second Admin.	118 63	4.92 4.25	2.11 1.77	29	.5657
5. School Conditions and Capacities	First Admin. Second Admin.	112 63	15.11 15.44	4.12 4.21	28	.4616
Total Instrument Score	First Admin. Second Admin.	101 58	51.20 50.38	11.41 11.81	23	.6147

for the instrument (S. M. Hord, personal communications, September 1996). It should be noted that Hord later changed the name of the instrument from AEL's field-test name to "School Professional Staff as Learning Community," which is more closely aligned to the purposes for which it was designed originally.

Concurrent Validity

The concurrent validity of the field test instrument was assessed by correlating its total score with that of an organizational climate instrument taken from Manning, Curtis, and McMillen (1996). Titled the "School Climate Questionnaire," this validity check instrument consisted of 10 items, many directly in line with items in the field test instrument. Again, a subsample of four high school faculties was involved in this portion of the field test. Although 190 of the school climate instruments were received for analysis, only 114 could be matched by their individual identification number that they were asked to write on both instruments. The correlation between the 17-item field test instrument and the 10-item school climate instrument for the 114 high school teachers in the subsample was .7489, which was significant at the .001 level.

Construct Validity

The two methods used to assess the construct validity of the instrument were the "known group" method and the factor analysis method. The results of each are described below.

Known group. Table 15 presents the t-test results for the known group school faculty versus the full group of teachers in the field test on the five dimension scores and the total instrument score. The number of teachers in the known group school was 18 or 19 on the six scores. For the full group of teachers, their numbers ranged from 595 on the total instrument score to 679 on dimension #2. The highest possible score on dimension #1 was 10 points and, with a mean score of 9.105, the known group was very close to the maximum, while the mean for the full group was 6.722. The 95% confidence interval for the difference in the mean was from almost 2 points to almost 3 points and the coefficient of variation for the full group was more than double that of the known group. The t-value was large at 10.52, after adjusting the degrees of freedom for unequal variances in the two groups. The difference between the known group and the full group on dimension #1 was significant at the .0001 level. The highest possible score on dimension #2 was 15 points and, with a mean of 14.222, the known group was very close to the maximum, while the mean for the full group was a little more than 3 points under that mean. The 95% confidence interval for the difference in the two means was from 2.84 to 3.65. The coefficient of variation value for the full group was more than four times larger than that of the known group. The t-value was a very large 16.38, after adjusting the degrees of freedom for the unequal variances in the two groups. The difference between the known group and the full group was significant at the .0001 level.

The highest possible score on dimension #3 was 25 points and the known group's mean of 23.00 was exactly two points below the maximum. The mean for the full group was 17.765 and its

Table 15

T-Test Results for the Known Group School Faculty Versus the Full Group of Teachers on the Five Dimensions and the Total Instrument Score

Group Name	Number of Teachers	Mean	Std. Dev.	Coeff. of Variation	95% Confidence Interval for Difference	t-Value	Degrees of Freedom	Signifi-cance Level
1. Principal's Facilitative Leadership								
Known Group	19	9.105	0.937	.103	1.91-2.85	10.52	22.23	.0001
Full Group	677	6.722	1.868	.278				
2. Shared Visions for Improvement								
Known Group	18	14.222	0.732	.051	2.84-3.65	16.38	29.60	.0001
Full Group	679	10.974	2.549	.232				
3. Collective Creativity and Learning								
Known Group	19	23.000	1.826	.079	4.31-6.16	11.76	22.93	.0001
Full Group	673	17.765	3.903	.220				
4. Classroom Observations and Feedback								
Known Group	19	7.368	1.342	.182	1.59-2.92	7.09	20.52	.0001
Full Group	649	5.113	2.042	.399				
5. School Conditions and Capacities								
Known Group	19	22.895	1.595	.070	4.28-6.47	14.09	25.85	.0001
Full Group	655	17.250	4.178	.241				
Total Instrument Score								
Known Group	18	76.944	4.452	.058	16.58-21.37	16.29	25.76	.0001
Full Group	595	59.973	12.327	.213				

standard deviation was more than double that of the known group. The 95% confidence interval for the two means was 4.31 to 6.16. The t-value was a large 11.76 after adjusting the degrees of freedom for the unequal variances in the two groups. This difference between the known group and the full group was significant at the .0001 level. Like the first dimension, the maximum possible score for dimension #4 was 10 points. With a mean score of 7.368 and a standard deviation of 1.342, the known group was considerably higher and had less dispersion than the full group. The 95% confidence interval for the two means was 1.59 to 2.92, which was a little wider interval than for the similarly-pointed first dimension. The resultant t-value of 7.09, although large, was the smallest in the table. Still, this difference between the known group and the full group was significant at the .0001 level. Like the third dimension, the maximum possible score for dimension #5 was 25 points. The mean scores were very similar to those obtained on the third dimension at 22.895 for the known group and 17.250 for the full group, as were the standard deviations and coefficients of variation. Likewise, at 4.28 to 6.47, the 95% confidence interval for the difference of the means is similar to the third dimension. The resulting t-value was a very large 14.09, with adjusted degrees of freedom of 25.85. As with all other dimensions, the difference in the mean scores between the known group and the full group was significant at the .0001 level.

Finally, Table 15 displays the t-test results for the total instrument score. Possible scores on the total instrument ranged from 17 to 85 points. The mean score for the known group was 76.944 and was 59.973 for the full group, yielding a very large difference of 16.971. Similarly, there was a very large difference in the standard deviations of 7.875, with the known groups being smaller. Also, there was a difference of .155 in the coefficient of variance, again with the known group's being smaller. The 95% confidence interval for the difference in the means was from 16.58 to 21.37. The resultant t-value was a very large 16.29, with adjusted degrees of freedom of 25.76. Consistent with all of the dimension scores, the difference in the mean scores for the total instrument between the known group and the full group was significant at the .0001 level.

Factor analysis. The second method for assessing the construct validity of the instrument was factor analysis. Exploratory factor analysis on the 595 cases with scores on all 17 items was completed in an attempt to locate and describe the number of factors within the items. The initial statistics from the principal axis factoring revealed that the communality values for the 17 items ranged from .49872 up to .67632. There was just one communality value in the .40s and ten in the .50s; the remaining six were in the .60s, indicating moderate shared variance among the items. The initial statistics showed two eigenvalues above 1.0 accounting for 59.0% of the variance in the scores. The largest eigenvalue was 8.7168, which accounted for 51.3% of the variance, while the second eigenvalue was 1.3061 and accounted for another 7.7% of the variance. The remaining 15 eigenvalues were under 1.0 and, collectively, they accounted for the remaining 41.0% of the variance.

The scree plot from the principal axis factoring displayed the largest eigenvalue at 8.717, followed by two at .944, four at .586, and the remaining ten at .195. Two factors were extracted in 12 iterations. The factor loadings for the first factor ranged from .56188 to .75728 and included all 17 items in the instrument. The second factor consisted of just two items (4a and 4b) with factor loadings of .54671 and .56471, respectively. Table 16 displays the communality values for the final statistics of the principal axis factoring in its second column. These final communality values ranged

Table 16

Final Communality Statistics and Factor Loadings from the Principal Axis Factoring and the Varimax and Oblique Rotations

Item Number	Final Commu-nality	Principal Axis		Varimax Rotation		Oblique Rotation	
		Factor I	Factor II	Factor I	Factor II	Factor I	Factor II
1a	.5481	.7078		.7161		.8015	
1b	.5268	.6922		.7036		.7890	
2a	.5466	.7110		.7114		.7924	
2b	.5802	.7387		.7270		.8037	
2c	.5945	.7573		.7204		.7838	
3a	.4766	.6866		.6219		.6604	
3b	.5737	.7471		.7010		.7578	
3c	.5128	.7129		.6422	.3169	.6800	
3d	.4748	.6860		.6177	.3054	.6539	
3e	.5200	.7209		.6040	.3941	.6146	
4a	.6505	.5929	.5467		.7768		.7736
4b	.6346	.5619	.5647		.7758		.7890
5a	.4965	.7044		.6086	.3552	.6302	
5b	.5051	.7088		.5757	.4166	.5752	
5c	.5578	.7466		.6238	.4107	.6339	
5d	.4452	.6399		.4451	.4971	.3992	.3539
5e	.5166	.6989		.5062	.5103	.4683	.3406

Notes: (1) Only loadings above .30 are displayed.
(2) In the oblique rotation, Factor I correlated with Factor II at .5684.

from .4452 to .6505, again an indication of moderate shared variation, but slightly lower than the initial communality values. For example, only two communality values were in the .60s, compared to six in the initial statistics. In the final statistics, the two factors accounted for 53.9% of the variance, with the larger factor accounting for 48.6% with an eigenvalue of 8.2538 and the smaller factor accounting for an additional 5.3% of the variance. However, the eigenvalue for the second factor was under 1.0 at .9070.

Table 16 displays the results from the principal axis factoring and both the unconstrained Varimax and oblique rotations of the factors. Factor loadings of less than .30 were omitted from the computer analysis printout and this display. From the principal axis factoring results in the third and fourth columns, data show that all 17 items loaded on Factor I. These loadings ranged from .5619 up to .7573. The two lowest loadings on Factor I were items 4a and 4b and these same two items were the only ones loading on Factor II at .5467 and .5647, respectively. In the unconstrained Varimax rotation, all of the items except 4a and 4b loaded on the first factor above .30. The Factor I loadings ranged from .4451 to .7270. Factor II in the Varimax rotation included all of the items from 3c to 5e at .30 or more. These Varimax Factor II loadings ranged from .3054 up to .7768. Items 4a and 4b had the two highest loadings at .7768 and .7758, respectively. The last two columns display the results of the unconstrained oblique rotation. Factor I in the oblique rotation included the exact same 15 items as the Varimax rotation—all but items 4a and 4b. Their factor loadings ranged from .3992 to .8037, slightly higher than those of the Varimax rotation. The second factor in the oblique rotation consisted of four items (4a, 4b, 5d, and 5e) with loadings of .7736, .7890, .3539, and .3406, respectively. On the oblique rotation, Factor I correlated with Factor II at .5684.

DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

Discussion and conclusions drawn from the findings are presented first in this section, followed by recommendations for use of the instrument and further research. These conclusions and recommendations are not presented in any particular order.

Discussion and Conclusions

The descriptive statistics from the field test of the new instrument were very encouraging. For example, the minimum and maximum scores were recorded for all 17 items and the standard deviations all were in the .91 to 1.20 range on the 5-point scale. The majority of the items were negatively skewed, but not by a lot. For the five dimensions and the total score, again all of the minimum and maximum scores were recorded and, while all of the mean scores were above the mid-points, four dimensions were negatively skewed. Even more encouraging were the descriptive statistics by the three levels of schools. At each level (elementary, middle/junior high, and high school), one or two schools consistently had higher scores on the five dimensions and the total score. If we define higher scores to be reflective of "more mature" school faculties, in terms of a professional learning community, then the descriptive results identify such "more mature" faculties involved in this field test. Conversely, if we define lower scores on the instrument to be reflective of "less mature" school faculties in terms of a professional learning community, then the descriptive results identify such "less mature" faculties at the elementary, middle/junior high, and high school levels.

Therefore, based on the descriptive statistics from the reformatted Hord instrument with 21 schools in the AEL Region, we conclude that it does differentiate among the schools on its five major dimensions and total score. Also, we conclude that the instrument differentiates the schools on its five dimensions and total score when the schools are grouped into the three levels of elementary, middle/junior high, and high school. The reformatted Hord instrument does measure and differentiate school faculties in terms of their "maturity" as professional learning communities.

The internal consistency (coefficient Alpha) reliabilities for the Hord instrument were very encouraging from the field test. All five dimension score coefficient Alphas were in the .80s, from .83 to .87, and the total instrument Alpha was .94. All but one of 17 items contributed to the total instrument Alpha value, as it would not have increased if any of those 16 items were deleted. The single exception was item 4b, but the Alpha reliability would increase only a trivial .0001 if it was deleted. This was a relatively short instrument; for example, two dimensions had only two items, one dimension had three items, and only two dimensions had five items. Yet, the Alpha reliability coefficients for the dimensions for the full group were very satisfactory and for the total instrument, excellent. Some of these high internal consistency reliabilities were due to the large sample size of 595 cases with complete data. However, these very satisfactory coefficient Alpha reliabilities held up rather well when the data were disaggregated by the 21 individual schools. Here, only one school (a small middle/junior high school) faculty produced an overall Alpha below .84. True, some Alphas

for individual dimensions were under .80, but there was no pattern to them, either by dimension or by school level, except that the second dimension produced lower reliabilities than the other four dimensions.

Therefore, we conclude that the reformatted Hord instrument yielded satisfactory internal consistency (coefficient Alpha) reliabilities for the five dimensions and the total instrument in this field test. These satisfactory coefficient Alpha relabilities were evident at both the full group and at the individual school level, although the second dimension had lower Alpha reliabilities than the other dimensions. There was no pattern in the Alpha reliabilities by the three levels of elementary, middle/junior high, and high school. Also, we conclude that the five dimensions were correlated with each other moderately and two (dimensions #3 and #5) were highly correlated (.90s) with the total instrument score.

In order to gain the cooperation of schools in this field test, they were assured that individual teacher names on their responses were not needed and the solicitation of a personal identification number (last four digits of Social Security number) was omitted at the printing stage. This proved to be a weakness in the field test as the lack of the solicitation of the personal identification number meant that this number had to be requested orally during the administration. One contact person volunteered to ask the teachers to supply this number and, while we are appreciative of the offer and the effort, the results proved to be very disappointing. This is most evident in the stability (test-retest) reliability part of the field test where only 30 of the high school teachers from three high schools in one state could be matched on any test-retest dimension scores. Further, only 23 of those teachers could be matched on the total instrument score. Given these less than ideal circumstances, the resulting coefficient of stability reliabilities for the five dimensions were moderate with values in the .40s and .50s. The coefficient of stability reliability for the total instrument score was higher at .61.

Thus, even though the coefficient of stability (test-retest) reliability value was computed on a smaller subsample than would be ideal, the resulting value for the total instrument score was marginally satisfactory, with the potential to increase if the sample size were to increase.

An extensive review of educational and business/corporate literature was completed by the original instrument developer. Hord developed items (descriptors and indicators) based on her four-year qualitative study of an elementary school operating as a professional learning community with four different principals over 15 years. In the second stage of the content validity assessment, three AEL staff carefully reviewed each descriptor and its set of related indicators for clarity, consistency, and intent. A few minor word changes were made by AEL staff, but not enough to change the intent in the original. AEL staff also reformatted the Hord instrument to make it easier to complete and to separate the descriptors graphically. In the third stage of content validity assessment, Hord reviewed AEL's minor changes and agreed with them.

Therefore, based on the three stages of the review of the items in the instrument, we conclude that the reformatted Hord instrument possesses sufficient content validity for its original intention of measuring the concept of a community of learners within the professional staff of K-12 schools.

A subsample of four high school faculties was involved in the concurrent validity aspect of this field-test study. Teachers in this part of the study completed a "School Climate Questionnaire" at the same time they completed the reformatted Hord instrument. The 10-item school climate instrument assessed many of the same qualities of the school as the target instrument. Of the four high school faculties in this concurrent validity check, 114 teachers could be matched by their personal identification numbers on both instruments. The correlation between the total scores of the two instruments was .75, significant at the .001 level.

Thus, with respect to the concurrent validity of the Hord instrument, we conclude that it does possess satisfactory correlation with the school climate instrument used in this field test with a subsample of four high school faculties in one state in AEL's Region.

This study benefitted by working closely with Dr. Shirley Hord, the developer of the original instrument. She agreed to administer the reformatted instrument to faculty of the school that she studied for several years and that she judged to be a "very mature" professional learning community. This school's staff became the "known group" in one method of assessing the construct validity of the instrument. The difference in scores between the "known group" and the full field-test group was very large and statistically significant (.0001 level) on all five dimensions and the total instrument score, favoring the "known group" in every comparison. For example, the difference in mean scores on the 17 to 85 point total score was 16.97.

Thus, regarding the construct validity of the instrument and its dimensions via the "known group" method, we conclude that the five dimensions and the total instrument do represent the constructs they purport to measure. Using the "known group" methodology, the reformatted instrument does appear to represent five distinct constructs of a professional learning community within schools and, when combined, they appear to represent the construct of a full professional learning community.

However, factor analysis, the second method of assessing construct validity, did not support the results of the "known group" method. It appears the best solution from the factor analysis data is a unitary factor that accounts for 54% of the variance. There is the beginning of a second factor in the two items in the fourth dimension of Classroom Observations and Feedback. But, in the end, the eigenvalue for this factor is less than 1.0 and its two items are a better fit with the other 15 items in one main construct than as a separate construct. The unidimensional nature of the 17 items from the factor analysis was surprising, given the earlier indications from the descriptive statistics and the internal consistency reliabilities that the five dimensions appeared to represent separate factors. Too, the "known group" construct validity results tended to support the notion of distinct constructs within each dimension. One reason for a unitary factor might be that the number of items for the dimensions—especially three of them—is small. Recall, two dimensions had only a pair of items and a third dimension had just three items. Another reason may be in the unique nature of the items with each one having three different indicators under the response scale, for a total of 51 full sentences the respondent had to read to complete the instrument. We have no empirical evidence to support this latter explanation, but experience leads us to believe that it is viable until proven otherwise.

Therefore, based on the factor analysis results, we conclude that the 17-item instrument represents a unitary construct of a professional learning community within schools. However, other results from this field test study lead us to conclude that there is the possibility of other factors existing within this larger construct, but we did not have enough evidence in this study to confirm their existence.

Hord developed the original version of the instrument to be a nominating device for her use in selecting research/study school sites to investigate over several years. Hord wanted to identify "more mature" school faculties in terms of being learning communities. Data from this field test of the reformatted instrument with 21 schools in AEL's Region confirms that it does differentiate professional staff by school at three levels, is very reliable internally, is marginally reliable in terms of its stability, possesses satisfactory concurrent validity with a school climate instrument, and displays very satisfactory construct validity when compared to a "known group" school faculty. Factor analysis results shows the present form of the instrument to represent a unitary factor within its 17 items.

Therefore, overall, we conclude that the present form of the 17-item instrument is very useful as a screening, filtering, or measuring device to assess the maturity of a school's professional staff as a learning community, especially when the total instrument score is used. We conclude that the five dimensions, as presently constructed, are useful for descriptive purposes in comparing different school faculties, but they do not possess sufficient evidence now to be labeled legitimately as factors or subscales.

Recommendations

The first recommendation from the field test of this instrument is to collect comparable data from more urban schools with more diverse teacher populations. This field test was conducted with 21 schools, divided reasonably across the three levels of elementary, middle/junior high, and high school. Also, these schools ranged in the size of their faculties from small to large. The number of teachers in the database (690) was very satisfactory for the field test of the instrument. However, only two of the schools (except for the one in the "known group") were located in an urban environment. We don't know if the inclusion of urban schools would have changed any of the instrument's data and/or results—and that is one of the weaknesses of this study. This study could have profited by including urban schools with faculties of more diverse cultural and ethnic backgrounds than represented by the 21 volunteer schools. We were fortunate to obtain as many cooperative schools as we did, and we appreciated the efforts of our contact persons in securing those schools. But hindsight is always better than foresight, and hindsight tells us we should have attempted more systematically to secure the cooperation of faculties in urban schools at all three levels.

The second recommendation is that the stability reliability and the concurrent validity analyses should be conducted with larger subsamples of teachers, especially teachers at the elementary and middle/junior high school levels. Part of the reason for the small subsamples for those two analyses

in this effort was a function of having to ask participating teachers to write in their personal identification numbers on all of the instruments, as this solicitation line was omitted in the field test. We are satisfied that we sought and obtained the cooperation of the contact person directly involved with administering the instruments to the four largest faculties in the group of 21 schools. These four faculties were high schools in one state. The recommendation here would be to conduct the stability reliability and concurrent validity analyses with elementary and middle/junior high school faculties.

A third recommendation is to include other instruments in the concurrent validity research with the reformatted Hord instrument. Only one instrument, a school climate questionnaire, was employed in the concurrent validity analysis. Although the results were satisfactory in this field test, there may be other instruments that also are related closely to the Hord instrument. Instrumentation in the area of professional learning communities in schools is in its infancy, but Louis and Marks (1996) and Cavanaugh and Dellar (1996) have developed new instruments measuring some of the same dimensions as the Hord instrument. A logical next step would be to use these other instruments in assessing the concurrent validity of the instrument developed by Hord. Related to the construct validity, future studies could include the use of qualitative data from school faculties that completed the instrument. This was not done as part of this field test.

Finally, our fourth recommendation concerns the next steps in the development of an instrument to measure the concept of a community of learners in the professional staff of K-12 schools. We concluded above that the reformatted, 17-item Hord instrument is a viable device for measuring the maturity of a school's professional learning community with the total instrument score. This field test of the instrument uncovered strong *possibilities* that there are distinct factors within the one dimension identified, but there is insufficient empirical evidence in this field test to support these other factors. We do know that the instrument used in this field test was unique in its construction. While the response scale of 1 to 5 was not unusual, the employment of three different, sentence-long indicators under each scale was unique. Our experience leads us to question how this very different response option layout may have affected how the teachers responded to the 17 items. We will never know from the data collected and analyzed in this study, but further development of this instrument may answer those questions. So, this recommendation calls for (1) the redesign of the instrument into a more typical format and response option and (2) then to conduct a pilot test and field test on this newly-revised version. Specifically, we recommend that a revised instrument be developed that addresses four points: (1) writing many more and shorter stem statements for each of the five dimensions; (2) redesigning the response options to be like those typically used in questionnaires, such as a 5- or 6-point Likert scale of Strongly Disagree to Strongly Agree; (3) reversing about one-third to one-half of the new item stems; and (4) switching the format from horizontal to vertical and typesetting the items in either one or two columns. If this is too costly or time-consuming, a less-expensive procedure would be to make the 51 indicator sentences individual item statements on the left side of a page and include a 5- or 6-point Likert response scale for each on the right side of the page. A related recommendation would be to solicit the respondent's personal identification number and basic demographic information—although the former is more crucial than the latter for field testing, since the unit of analysis is the school, and the identification number is essential for matching instruments.

REFERENCES

Appalachia Educational Laboratory. (1995). *AEL technical proposal for operation of the regional educational laboratory in the Appalachian region for 1996-2000*. Charleston, WV: Author.

Appalachia Educational Laboratory. (1997, April). *Quarterly report: January 1, 1997, through March 31, 1997*. Charleston, WV: Author.

Block, P. (1993). *Stewardship: Choosing service over self-interest*. San Francisco: Berrett-Koehler.

Boyd, V., & Hord, S. M. (1994a, April). *Principals and the new paradigm: Schools as learning communities*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Boyd, V. & Hord, S. M. (1994b). *Schools as learning communities. Issues...About change, 4*(1). Austin, TX: Southwest Educational Development Laboratory.

Cavanaugh, R. F. & Dellar, G. B. (1996, March). *The development of an instrument for investigating school culture*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Deal, T. E. & Kennedy, A. A. (1982). *Corporate cultures*. Reading, MA: Addison-Wesley Publishing.

Galagan, P. (1994, December). Reinventing the profession. *Training & Development, 48(*12), 20-27.

Hord, S. M. (1996a). *Communities of continuous inquiry and improvement (Draft Copy)*. Austin, TX: Southwest Educational Development Laboratory.

Hord, S. M. (1996b). *Descriptors of professional learning communities*. Austin, TX: Southwest Educational Development Laboratory.

Hord, S. M. & Boyd, V. (1995, Winter). Staff development fuels a culture of continuous improvement. *Journal of Staff Development, 16* (1), 10-15.

Louis, K. S. & Marks, H. (1996, March). *Does professional community affect the classroom? Teacher's work and student experiences in restructuring schools*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Manning, G., Curtis, K., & McMillen, S. (1996). *Building community: The human side of work.* Cincinnati, OH: Thomas Executive Press.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Department of Education, U.S. Government Printing Office.

Neuman, F. M. & Associates (1996). *Authentic achievement: Restructuring schools for intellectual quality*. San Francisco: Jossey-Bass.

Plank, D. (1997, April). *Dreams of community*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Quality Education Data. (1996a). *QED state school guide: Kentucky, 1996-1997* (17th ed.). Denver, CO: Author.

Quality Education Data. (1996b). *QED state school guide: Tennessee, 1996-1997* (17th ed.). Denver, CO: Author.

Quality Education Data. (1996c). *QED state school guide: Virginia, 1996-1997* (17th ed.). Denver, CO: Author.

Quality Education Data. (1996d). *QED state school guide: West Virginia, 1996-1997* (17th ed.). Denver, CO: Author.

Senge, P. (1990). *The fifth discipline: The art and practice of the organization*. New York: Currency Doubleday.

Whyte, D. (1994). *The heart aroused: Poetry and the preservation of the soul in corporate America*. New York: Currency Doubleday.

APPENDICES

APPENDIX A:

Hord's First Version Professional Learning Comunity Instrument

Nomination for School Research/Study Site

_____
Your Name

_____
Your phone number

_____
The nominated school's name

_____
Grade levels in the school

_____
School's phone number

_____
Principal's name

_____/_____
# of Pupils # of Professional Staff

How long have you known this school?

_____

What is your relationship to the school?

_____

Descriptors of Professional Learning Communities

Please consider where you believe the nominated school is in its development of each of the five Roman-numbered descriptors (below). Each descriptor has a series of five-point scale indicators that you may mark to show your assessment. You may mark any numeral on the scale, in order to express the degree to which the school has developed.

I. School administrator(s) participate democratically with teachers sharing power, authority, and decision making.

5	4	3	2	1

Although there are some legal and fiscal decisions required of the principal, this person consistently involves the staff in discussing and making participative decisions about most of the school issues.	Administrator(s) invites advice and counsel from the staff and then makes decisions themself(ves).	Administrator(s) never shares information with the staff nor provides opportunity to be involved in decision making.

5	4	3	2	1

Administrator(s) involves the entire staff.	Administrator(s) involves a small committee, council or team of staff.	Administrator(s) does not involve any staff.

II. The development of shared visions for improvement have an undeviating focus on student learning, and are consistently referenced for the staff's work.

5	4	3	2	1

Visions for improvement are discussed by the entire staff such that consensus and a vision shared by all results.	Visions for improvement are not thoroughly explored, with some staff agreeing while others do not.	Visions for improvement held by the staff are widely divergent and with little agreement.

5	4	3	2	1
Visions for improvement are always focused on students and teaching and learning.		Visions for improvement are sometimes focused on students and teaching and learning.		Visions for improvement do not target students and teaching and learning.

5	4	3	2	1
Visions for improvement target high quality learning experiences for all students.		Visions for improvement address quality learning experiences in terms of students' abilities.		Visions for improvement do not include concerns about the quality of learning experiences.

III. Staff's collective learning and application of the learnings (taking action) create high intellectual learning tasks and solutions to address student needs.

5	4	3	2	1
The entire staff meets to discuss issues, share information, and learn with and from each other.		Subgroups of the staff meet to discuss issues, share information, and learn with and from each other.		Individuals discuss issues, share information, and learn with and from each other.

5	4	3	2	1
The staff meets regularly and frequently on substantive student-centered educational issues.		The staff meets occasionally on substantive student-centered educational issues.		The staff never meets to consider substantive educational issues.

5	4	3	2	1
The staff meets to discuss the quality of their teaching and students' learning.		The staff do not question their instructional practices nor its influence on student learning.		The staff discusses fund non-teaching and learning issues.

5	4	3	2	1

The staff, based on their learnings, makes and implements plans that address students' needs, more effective teaching, and more successful student learning.	The staff occasionally acts on their learnings and makes and implements plans to improve teaching and learning.	The staff does not act on their learning.

5	4	3	2	1

The staff debriefs and assesses the impact of their actions and makes revisions.	The staff infrequently assesses their actions and seldom makes revisions based on the results.	The staff does not assess their work.

IV. Peers review and give feedback based on observing each other's classroom behaviors in order to increase individual and organizational capacity.

5	4	3	2	1

Staff regularly and frequently visit and observe each other's classroom teaching.	Staff occasionally visit and observe each other's teaching.	Staff never visit their peers' classrooms.

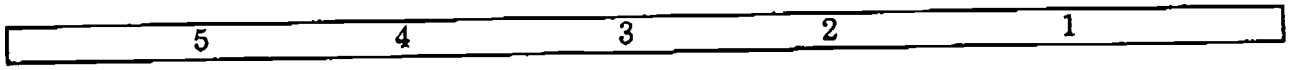
5	4	3	2	1

Staff provide feedback to each other about teaching and learning based on their observations.	Staff discuss non-teaching issues after observations.	Staff do not interact after observations.

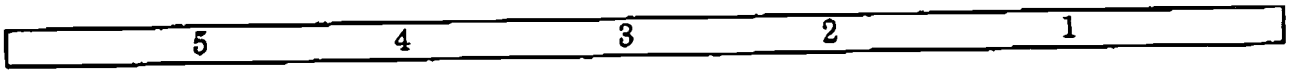
V. Conditions and capacities support the school's arrangement as a professional learning organization.

5	4	3	2	1

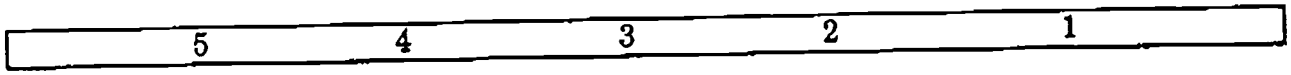
Time is arranged and committed for whole staff interactions.	Time is arranged but frequently the staff fails to meet.	Staff cannot arrange time for interacting.



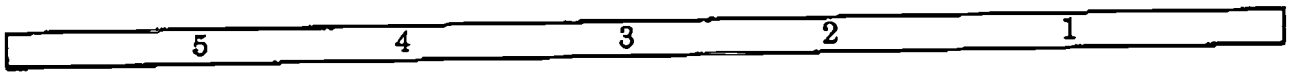
The site, structure, and arrangement of the school facilitate staff proximity and interaction.	While the facility and school membership are large, the staff are working to maximize existing arrangements for interaction.	The staff takes no action to manage the facility and personnel for interaction.



A variety of processes and procedures are used to encourage staff communication.	A single communication device exists and is sometimes used to share information.	Communication devices are not given attention.



Trust and openess to feedback and learning characterize all the staff.	Some of the staff are trusting and open to learning from other staff.	Trust does not exist among the staff.



Caring, collaborative, and productive relationships exist among all the staff.	Caring and collaboration are inconsistently demonstrated among the staff.	Staff are isolated and work alone at their tasks.

APPENDIX B:

Hord's Instrument Reformatted by AEL Staff

School as Learning Organization*

Directions: This questionnaire concerns your perceptions about your school as a learning organization. There are no right or wrong responses. Please consider where you believe your school is in its development of each of the five numbered descriptors shown in bold-faced type on the left. Each sub-item has a five-point scale. On each scale, circle the number that best represents the degree to which you feel your school has developed.

Last Four Social Security Numbers: ___ ___ ___ ___

Date: _____

School: _____

1. School administrators participate democratically with teachers sharing power, authority, and decision making.

1a

5	4	3	2	1
Although there are some legal and fiscal decisions required of the principal, school administrators consistently involve the staff in discussing and making decisions about most school issues.		Administrators invite advice and counsel from the staff and then make decisions themselves.		Administrators never share information with the staff nor provide opportunities to be involved in decision making.

1b

5	4	3	2	1
Administrators involve the entire staff.		Administrators involve a small committee, council, or team of staff.		Administrators do not involve any staff.

---

2. Shared visions for school improvement have an undeviating focus on student learning and are consistently referenced for the staff's work.

2a

5	4	3	2	1
Visions for improvement are discussed by the entire staff such that consensus and a shared vision results.		Visions for improvement are not thoroughly explored; some staff agree and others do not.		Visions for improvement held by the staff are widely divergent.

2b

5	4	3	2	1
Visions for improvement are always focused on students and teaching and learning.		Visions for improvement are sometimes focused on students and teaching and learning.		Visions for improvement do not target students and teaching and learning.

2c

5	4	3	2	1
Visions for improvement target high quality learning experiences for all students.		Visions for improvement address quality learning experiences in terms of students' abilities.		Visions for improvement do not include concerns about the equality of learning experiences.

Staff's collective learning and application of the learnings (taking action) create high intellectual learning tasks and solutions to address student needs.

3a

5	4	3	2	1
The entire staff meets to discuss issues, share information, and learn with and from each other.		Subgroups of the staff meet to discuss issues, share information, and learn with and from each other.		Individuals discuss issues, share information, and learn with and from each other.

3b

5	4	3	2	1
The staff meets regularly and frequently on substantive student-centered educational issues.		The staff meets occasionally on substantive student-centered educational issues.		The staff never meets to consider substantive educational issues.

3c

5	4	3	2	1
The staff discusses the quality of their teaching and students' learning.		The staff does not often discuss their instructional practices nor its influence on student learning.		The staff basically discusses non-teaching and non-learning issues.

3d

5	4	3	2	1
The staff, based on their learnings, makes and implements plans that address students' needs, more effective teaching, and more successful student learning.		The staff occasionally acts on their learnings and makes and implements plans to improve teaching and learning.		The staff does not act on their learning.

3e

5	4	3	2	1
The staff debriefs and assesses the impact of their actions and makes revisions.		The staff infrequently assesses their actions and seldom makes revisions based on the results.		The staff does not assess their work.

---

4. Peers review and give feedback based on observing each other's classroom behaviors in order to increase individual and organizational capacity.

4a

5	4	3	2	1
Staff regularly and frequently visit and observe each other's classroom teaching.		Staff occasionally visit and observe each other's teaching.		Staff never visit their peers' classrooms.

4b

5	4	3	2	1
Staff provide feedback to each other about teaching and learning based on their classroom observations.		Staff discuss non-teaching issues after classroom observations.		Staff do not interact after classroom observations.

Conditions and capacities support the school's arrangement as a professional learning organization.

5a

5	4	3	2	1
Time is arranged and committed for whole staff interactions.		Time is arranged but frequently the staff fails to meet.		Staff cannot arrange time for interacting.

5b

5	4	3	2	1
The site, structure, and arrangements of the school facilitate staff proximity and interaction.		While the facility and school membership are large, the staff are working to maximize existing arrangements for interaction.		The staff takes no action to manage the facility and personnel for interaction.

5c

5	4	3	2	1
A variety of processes and procedures are used to encourage staff communication.		A single communication exists and is sometimes used to share information.		Communication devices are not given attention.

5d

5	4	3	2	1
Trust and openness characterize all the staff.		Some of the staff are trusting and open.		Trust and openness do not exist among the staff.

5e

5	4	3	2	1
Caring, collaborative, and productive relationships exist among all the staff.		Caring and collaboration are inconsistently demonstrated among the staff.		Staff are isolated and work alone at their task.

*Hord, Shirley M. (1996). Austin, TX: Southwest Educational Development Laboratory.

APPENDIX C:

Completed Evaluation *Standards* Checklist

Citation Form

The Program Evaluation Standards (1994, Sage) guided the development of this (check one):

_____ request for evaluation plan/design/proposal
_____ evaluation plan/design/proposal
_____ evaluation contract
_____ evaluation report
XX other: Field Test of Instrument

To interpret the information provided on this form, the reader needs to refer to the full text of the standards as they appear in Joint Committee on Standards for Educational Evaluation, *The Program Evaluation Standards* (1994), Thousand Oaks, CA, Sage.

The *Standards* were consulted and used as indicated in the table below (check as appropriate):

Descriptor		The Standard was deemed applicable and to the extent feasible was taken into account.	The Standard was deemed applicable but could not be taken into account.	The Standard was not deemed applicable.	Exception was taken to the Standard.
U1	Stakeholder Identification				
U2	Evaluator Credibility	XXXX			
U3	Information Scope and Selection	XXXX			
U4	Values Identification	XXXX			
U5	Report Clarity	XXXX			
U6	Report Timeliness and Dissemination	XXXX			
U7	Evaluation Impact	XXXX			
F1	Practical Procedures	XXXX			
F2	Political Viability			XXXX	
F3	Cost Effectiveness	XXXX			
P1	Service Orientation	XXXX			
P2	Formal Agreements			XXXX	
P3	Rights of Human Subjects	XXXX			
P4	Human Interactions	XXXX			
P5	Complete and Fair Assessment	XXXX			
P6	Disclosure of Findings	XXXX			
P7	Conflict of Interest	XXXX			
P8	Fiscal Responsibility	XXXX			
A1	Program Documentation	XXXX			
A2	Context Analysis	XXXX			
A3	Described Purposes and Procedures	XXXX			
A4	Defensible Information Sources	XXXX			
A5	Valid Information	XXXX			
A6	Reliable Information	XXXX			
A7	Systematic Information	XXXX			
A8	Analysis of Quantitative Information	XXXX			
A9	Analysis of Qualitative Information	XXX (only content validity part)			
A10	Justified Conclusions	XXXX			
A11	Impartial Reporting	XXXX			
A12	Metaevaluation	XXXX			

Name (typed) Merrill L. Meehan Date: July 9, 1997

(signature)

Position or Title: Senior Research and Evaluation Specialist

Agency: Appalachia Educational Laboratory

Address: P. O. Box 1348, Charleston, WV 25325

Relation to Document: Lead Author
(e.g., author of document, evaluation team leader, external auditor, internal auditor)

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

TM030028

NOTICE

REPRODUCTION BASIS

☑ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").